

# ON NON-LOCAL VARIATIONAL PROBLEMS WITH LACK OF COMPACTNESS RELATED TO NON-LINEAR OPTICS

DIRK HUNDERTMARK AND YOUNG-RAN LEE

**ABSTRACT.** We give a simple proof of existence of solutions of the dispersion management and diffraction management equations for zero average dispersion, respectively diffraction. These solutions are found as maximizers of non-linear and non-local variational problems which are invariant under a large non-compact group. Our proof of existence of maximizer is rather direct and avoids the use of Lions' concentration compactness argument or Ekeland's variational principle.

## 1. INTRODUCTION

**1.1. The variational problems.** In this paper we are concerned with the existence of maximizers for two non-local non-linear variational problems,

$$P_\lambda^c := \sup \left( \mathcal{Q}_\mu^c(f, f, f, f) \mid f \in L^2(\mathbb{R}), \|f\|_{L^2}^2 = \lambda \right), \quad (1.1)$$

respectively,

$$P_\lambda^d := \sup \left( \mathcal{Q}_\mu^d(f, f, f, f) \mid f \in l^2(\mathbb{Z}), \|f\|_{l^2}^2 = \lambda \right), \quad (1.2)$$

for  $\lambda > 0$ . Here the four-linear functionals  $\mathcal{Q}_\mu^{c/d}$  are given by

$$\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) := \int_{\mathbb{R}} \int_{\mathbb{R}} \overline{(T_r f_1)(x)} (T_r f_2)(x) \overline{(T_r f_3)(x)} (T_r f_4)(x) dx \mu(dr) \quad (1.3)$$

in the continuous case, where  $\mu$  is a suitable measure on  $\mathbb{R}$ , the operator  $T_r := e^{ir\partial_x^2}$  is the unitary solution operator for the free Schrödinger equation in one dimension, and  $f_j \in L^2(\mathbb{R})$  for  $j = 1, 2, 3, 4$ , respectively

$$\mathcal{Q}_\mu^d(f_1, f_2, f_3, f_4) := \sum_{x \in \mathbb{Z}} \int_{\mathbb{R}} \overline{(S_r f_1)(x)} (S_r f_2)(x) \overline{(S_r f_3)(x)} (S_r f_4)(x) \mu(dr) \quad (1.4)$$

in the discrete case, where, with  $\Delta$  the discrete Laplacian given by  $\Delta f(x) = f(x+1) + f(x-1) - 2f(x)$  we denote by  $S_r := e^{ir\Delta}$  the solution operator of the free discrete Schrödinger equation in one dimension and  $f_j \in l^2(\mathbb{Z})$ ,  $j = 1, 2, 3, 4$ .

For the existence of maximizers of (1.2) we only need that  $\mu$  is a bounded measure with bounded support, see Theorem 1.2 and for the existence of maximizers of (1.1) we need that the measure  $\mu$  has a density  $\psi$  lying in suitable  $L^p$ -spaces, see Theorem 1.1.

---

©2009 by the authors. Faithful reproduction of this article, in its entirety, by any means is permitted for non-commercial purposes.

Supported in part by NSF-grant DMS-0803120 (D.H.) and the National Research Foundation of Korea (NRF)-grant 2009-0064945 (Y.-R.L.).

Our interest in these variational problems stems from the fact that the maximizers of (1.1), respectively (1.2), yield (quasi-)periodic breather type solutions, the dispersion management solitons, of the dispersion managed non-linear Schrödinger equation, respectively the diffraction management solitons for the diffraction managed discrete non-linear Schrödinger equation. The dispersion management solitons have attracted a lot of interest in the development of ultra-fast long-haul optical data transmission fibers and the diffraction management solitons were studied in some new discrete waveguide array designs. We address the connection of the two variational problems above with non-linear optics later in section 1.2.

The standard approach to show the existence of a maximizer of (1.1), respectively, (1.2), is to identify it as the strong limit of a suitable maximizing sequence, i.e., in the continuous case a sequence  $(f_n)_{n \in \mathbb{N}} \subset L^2(\mathbb{R})$  with  $\|f_n\|_2^2 = \lambda$  and  $P_\lambda^c = \lim_{n \rightarrow \infty} \mathcal{Q}_\mu^c(f_n, f_n, f_n, f_n)$ . The problem is that the two variational problems above are invariant under translations of  $L^2(\mathbb{R})$ , respectively  $l^2(\mathbb{Z})$ . The invariance of  $(l^2(\mathbb{Z}))^4 \ni (f_1, f_2, f_3, f_4) \mapsto \mathcal{Q}_\mu^d(f_1, f_2, f_3, f_4)$  under simultaneous translations of the  $f_j$  follows simply from the invariance of the discrete Laplacian  $\Delta$  and hence of  $S_r = e^{ir\Delta}$  under shifts. Even worse, due to the Galilei invariance of the free Schrödinger evolution the maximization problem (1.1) is invariant under translations *and* boosts, i.e., translation in momentum space, of  $L^2(\mathbb{R})$  for every sensible choice of the measure  $\mu$ , see the discussion in Appendix C. Thus the two variational problems above are invariant under a large non-compact group of transformations leading to a *loss of compactness*: maximizing sequences can easily converge weakly to zero.

Under the assumption that the measure  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , that is, the uniform distribution on  $[0,1]$ , this loss of compactness in the variational problem (1.1) was overcome by Kunze in [30], who used a very tricky application of the Lions' concentration compactness principle, [37], first in Fourier-space and then in real space, to compensate for the loss of compactness due to shifts and boosts. The existence of maximizers of (1.2), again under the same assumption that  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , was shown by Stanislavova in [49], using Ekeland's variational principle [17, 18, 27].

In this paper we give an alternative approach very much different from [30] and [49] for the existence of maximizers for the variational problems (1.1) and (1.2) which we believe is not only very natural but has several additional advantages:

- 1) We show existence of maximizers of (1.1) under *very weak* conditions on the measure  $\mu$ , see Theorem 1.1, and the existence of maximizers of (1.2) under the *weakest possible* assumption on  $\mu$ , see Theorem 1.2 and Remark 1.3.i.
- 2) Our approach avoids the use of 'heavy machinery' like Lions' concentration compactness or Ekeland's variational principle from the calculus of variations and gives easily more information about maximizing sequences. Whereas [30] and [49] show that there exists at least one suitable maximizing sequence which is strongly converging, we show that the loss of compactness is much milder than one would naively expect: *Any* maximizing sequence for (1.1), respectively (1.2), is tight, i.e., stays in a compact subset of  $L^2(\mathbb{R})$ , respectively  $l^2(\mathbb{Z})$ , modulo translations and boosts in  $L^2$ , respectively modulo translations in  $l^2$ , see Propositions 2.4 and 2.5.

- 3) To conclude that a maximizer of (1.1), respectively (1.2), exists, we use a simple characterization of strong convergence in  $L^2$ , respectively  $l^2$ , in terms of ‘weak convergence’ and ‘tightness’. This is done in Lemma A.1 and A.4 whose proof is rather straightforward and uses only some simple properties of compact operators.
- 4) We believe that these ideas will be useful in the study of other variational problems on  $L^2$ , respectively  $l^2$ .

Our main results concerning the variational problems (1.1) and (1.2) are

**Theorem 1.1** (Existence, continuous case). *Assume that the measure  $\mu$  has a density  $\psi$  with  $\psi \in L^2(\mathbb{R})$ . Then the variational problem (1.1) is well-posed, i.e.,  $P_\lambda^c < \infty$  for all  $\lambda > 0$ . Moreover, if the density  $\psi \in L^2(\mathbb{R}) \cap L^4(\mathbb{R}) \cap L^4(\mathbb{R}, t^2 dt)$ , then for any  $\lambda > 0$ , there exists a maximizer for the variational problem (1.1), i.e., there exists  $f \in L^2(\mathbb{R})$ ,  $\|f\|_2^2 = \lambda$ , such that*

$$\mathcal{Q}_\mu^c(f, f, f, f) = \sup \left( \mathcal{Q}_\mu^c(g, g, g, g) \mid g \in L^2(\mathbb{R}), \|g\|_2^2 = \lambda \right).$$

*This maximizer is also a solution of the dispersion management equation (1.22) for some Lagrange multiplier  $\omega > 0$ .*

In the discrete case we have an existence result under the ‘weakest possible’ assumption on the measure  $\mu$ .

**Theorem 1.2** (Existence, discrete case). *The variational problem (1.2) is well posed if  $\mu$  is a bounded measure. If, in addition, the measure  $\mu$  has bounded support, then for any  $\lambda > 0$ , there exists a maximizer for the variational problem (1.2). This maximizer is also a solution of the diffraction management equation (1.23) for some Lagrange multiplier  $\omega > 0$ .*

**Remarks 1.3.** (i) As we will see in section 1.2 the requirement that  $\mu$  is a probability measure with bounded support arises naturally in the study of diffraction management solitons. In this sense, the condition on  $\mu$  in Theorem 1.2 is optimal and the assumptions on the density of  $\mu$  in Theorem 1.1 are not very restrictive since in this case  $\psi \in L^1(\mathbb{R})$  with compact support and thus, by interpolation, the assumptions in Theorem 1.1 reduce to the additional requirement that  $\psi \in L^4(\mathbb{R})$ . In particular, the density  $\psi$  can have some strong local singularities which, via (1.14) below, yields existence of breather type solutions in dispersion managed glass fiber cables under mild conditions on the dispersion profile  $d_0$  of the fiber. For example, any locally continuous dispersion profile  $d_0$  which is bounded away from zero is allowed,  $d_0$  can even have (isolated) zeros, as long as they are approached slowly enough.

(ii) The existence Theorem 1.1 gives no further information about the regularity of the maximizer. If  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , Kunze’s existence proof, [30], shows that the maximizer is bounded. In [48] Stanislavova then showed that Kunze’s maximizer is infinitely often differentiable. Only recently it was shown in [24] that any weak solution  $f \in L^2(\mathbb{R})$  of the dispersion management equation (1.22) is a Schwartz function, i.e., it is infinitely often differentiable and all its derivatives decay faster than algebraically at infinity. All these result so far need that  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , but, as we shall see, the regularity result of [24] easily carries over to all  $\mu$  considered in Theorem 1.1, see Remark B.6.ii.

(iii) Again if  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , Lushnikov gave convincing but non-rigorous arguments in [38] that the maximizer of (1.1) should have the asymptotic form

$$f(x) \sim A \cos(a_2 x^2 + a_1 x + a_0) e^{-b|x|} \quad \text{as } |x| \rightarrow \infty$$

for some  $a_j$  and  $b > 0$ . In particular, the maximizer should be exponentially decaying. In [19] we show that any solution  $f \in L^2(\mathbb{R})$  of the dispersion management equation (1.22), so also any maximizer of (1.1), together with its Fourier transform is exponentially decaying if  $\mu$  has density  $\mathbf{1}_{[0,1]}$ , confirming part of Lushnikov's conjecture. In particular, even though it is no longer an elliptic equation, the singular limit (1.22) of the dispersion management equation still enjoys very strong regularity properties: any solution of it is analytic in a strip containing the real line under suitable conditions on the density of  $\mu$ .

(iv) As already mentioned, the existence of maximizers of the discrete maximization problem (1.2) was shown in [49] if  $\mu$  has density  $\mathbf{1}_{[0,1]}$ . Moreover, [49] shows that in this case the maximizer decays faster than algebraically at infinity. In [25] we show that under the conditions of the Existence Theorem 1.2 every maximizer and, more generally, any solution of the discrete Gabitov–Turitsyn equation (1.23) for vanishing average diffraction, is even *super-exponentially* decaying. More precisely, the bound

$$\limsup_{|x| \rightarrow \infty} ((|x| + 1) \ln(|x| + 1))^{-1} \ln |f(x)| \leq -\frac{1}{4} \quad (1.5)$$

holds for any solution  $f \in l^2(\mathbb{Z})$  of (1.23) if  $\mu$  is a bounded measure with compact support. Thus, unlike in the continuous case, we have the super-exponential decay estimate (1.5) for any solution under the same condition on the measure  $\mu$  as needed for existence.

Our paper is organized as follows: In the next section we discuss how the maximization problems (1.1), respectively (1.2), arise naturally in problems in non-linear optics where the dispersion, respectively diffraction, is strongly periodically varied. In section 2 we derive our main tools for the existence of maximizers for the variational problems (1.1) and (1.2). In Proposition 2.4 we show that any maximizing sequence for (1.1) can be shifted and boosted, so that it is tight both in real and Fourier space. A similar result, Proposition 2.5, is shown in the discrete case. These two propositions are the key for avoiding the use of Lions' concentration compactness argument or Ekeland's variational principle. The core of the argument is given in Lemmata 2.1, respectively, Lemma 2.2, which show that a near maximizer cannot break up in real and Fourier space. Strong convergence of suitably translated and boosted, respectively translated, maximizing sequences then follows from a simple characterization of strong convergence in  $L^2$ , respectively  $l^2$ , in Lemma A.1 and A.4. The proof of Lemma 2.1 is based on multi-linear refinement of the Strichartz inequality in real and Fourier space. These refinements are extensions of the multi-linear estimates developed in [24] and discussed in Appendix B.1. In the discrete case, we use the multi-linear estimates developed in [25], see Appendix B.2.

**1.2. The connection with non-linear optics.** Our main motivation for studying (1.1), respectively (1.2), comes from the fact that the maximizers of these two variational problems are related to breather-type solutions of the dispersion managed non-linear Schrödinger equation

$$i\partial_t u = -d(t)\partial_x^2 u - |u|^2 u, \quad (1.6)$$

respectively its discrete version,

$$i\partial_t u(t, x) = -d(t)(\Delta u)(t, x) - |u(t, x)|^2 u(t, x), \quad x \in \mathbb{Z}, \quad (1.7)$$

where the dispersion/diffraction  $d(t)$  is parametrically modulated. The continuous non-linear Schrödinger equation (1.6), respectively its discrete version (1.7), describe a wide range of different physical phenomena in such diverse areas as solid states physics, some biological systems, Bose-Einstein condensation, and continuous and discrete non-linear optics, e.g., glass-fiber cables and optical waveguide arrays, see, e.g., [7, 14, 46, 52, 54].

In non-linear optics (1.6) describes the evolution of a pulse in a frame moving with the group velocity of the signal through a glass fiber cable, see [53]. As a *warning*: with our choice of notation the variable  $t$  denotes the position along the glass fiber cable and  $x$  the (retarded) time. Hence  $d(t)$  is *not varying in time* but denotes indeed a dispersion *varying along* the optical cable. The discrete version (1.9) describes an array of wave-guides where  $t$  is the distance along the waveguide, the now discrete variable  $x \in \mathbb{Z}$  denotes the location of an array element, and  $d(t)$  the total diffraction along the waveguide. For physical reasons it is not a restriction to assume that  $d$  is piecewise constant, but we will not make this assumption in this paper.

In the continuous case the dispersion management idea, i.e., the possibility to periodically manage the dispersion by put alternating sections with positive and negative dispersion together in an optical glass-fiber cable to compensate for dispersion of the signal was predicted by Lin, Kogelnik, and Cohen already in 1980, see [36], and then implemented by Chraplyvy and Tkach for which they received the Marconi prize in 2009. The periodically varying dispersion creates a new optical fiber type enabling the development of long-haul optical fiber transmission systems with record breaking capacities beyond one Terabit/second per fiber which equates to a 100-fold capacity increase in the last ten years, [1, 10, 11, 21, 22, 29, 32, 33, 36, 39, 41, 42]. Thus dispersion management technology has been of fundamental importance for ultra-high speed data transfer through glass fiber cables over intercontinental distances and is now widely used commercially. For a review see [55, 56]. Discrete solitons in an optical waveguide array, on the other hand, were theoretically predicted in [12]. Nearly a decade later they were experimentally studied, [15], and as in the continuous case localized stable non-linear waves were found. Recently a zigzag diffraction management geometry in discrete optical waveguides was proposed in [16] in order to create low power stable discrete pulses which can be more easily observed experimentally.

In both cases, the periodic modulation of the dispersion, respectively diffraction, can be described by the ansatz

$$d(t) = \varepsilon^{-1} d_0(t/\varepsilon) + d_{av}. \quad (1.8)$$

Here  $d_{av} \geq 0$  is the average component and  $d_0$  its mean zero part which, by scaling, we can assume to have period two. For small  $\varepsilon$  the equation (1.8) describes a fast strongly

varying dispersion, respectively diffraction, which corresponds to the regime of *strong* dispersion, respectively diffraction, management.

Since (1.6) and (1.7) are formally very similar we will combine them into the equation

$$i\partial_t u = d(t)Au - |u|^2 u \quad (1.9)$$

on the Hilbert space  $X$  where we call the choice  $X = L^2(\mathbb{R})$  and  $A = -\partial_x^2 = -\frac{\partial^2}{\partial x^2}$  the continuous case and the discrete case is given by  $X = l^2(\mathbb{Z})$  and  $A = -\Delta$  the discrete Laplacian. We seek to rewrite (1.9) into a more amenable form in order to find breather type solutions. Let  $D(t) = \int_{-1}^t d_0(s) ds$  and note that as long as  $d_0$  is locally integrable and has period two with mean zero,  $D$  is also periodic with period two. Furthermore,  $U_r = e^{-irA}$  is a unitary operator and thus the unitary family  $t \mapsto U_{D(t/\varepsilon)}$  is periodic with period  $2\varepsilon$ . Making the ansatz  $u(t, x) = (U_{D(t/\varepsilon)}v(t, \cdot))(x)$  in (1.9), a short calculation shows

$$i\partial_t v = d_{\text{av}}Av - U_{D(t/\varepsilon)}^{-1} [|U_{D(t/\varepsilon)}v|^2 U_{D(t/\varepsilon)}v] \quad (1.10)$$

which is equivalent to (1.9) and still a non-autonomous equation.

For small  $\varepsilon$ , that is, in the regime of strong dispersion/diffraction management,  $U_{D(t/\varepsilon)}$  is fast oscillating in the variable  $t$ , hence the solution  $v$  should evolve on two widely separated time-scales, a slowly evolving part  $v_{\text{slow}}$  and a fast, oscillating part which is hopefully small. Analogously to Kapitza's treatment of the unstable pendulum which is stabilized by fast oscillations of the pivot, see [34], the effective equation for the slow part  $v_{\text{slow}}$  was derived by Gabitov and Turitsyn [21, 22] in the continuous case and in [2, 3, 4] in the discrete case. It is given by integrating the fast oscillating term containing  $U_{D(t/\varepsilon)}$  over one period in  $t$ ,

$$\begin{aligned} i\partial_t v_{\text{slow}} &= d_{\text{av}}Av_{\text{slow}} - \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} U_{D(r/\varepsilon)}^{-1} [|U_{D(r/\varepsilon)}v_{\text{slow}}|^2 U_{D(r/\varepsilon)}v_{\text{slow}}] dr \\ &= d_{\text{av}}Av_{\text{slow}} - \frac{1}{2} \int_{-1}^1 U_{D(r)}^{-1} [|U_{D(r)}v_{\text{slow}}|^2 U_{D(r)}v_{\text{slow}}] dr. \end{aligned} \quad (1.11)$$

This averaging procedure leading to (1.11) was rigorously justified for the profile  $d_0 = \mathbf{1}_{[-1,0)} - \mathbf{1}_{[0,1)}$ , in [59] in the continuous case and in [40] and [44] in the discrete case: given an initial condition  $f$ , the solutions of (1.11) and (1.10) stay  $\varepsilon$  close – measured in suitable Sobolev norms – over long distances  $0 \leq t \leq C/\varepsilon$ , see [59] and [40, 44] for the precise formulation. Thus of special interest are stationary solutions of (1.11), which can be found making the ansatz

$$v_{\text{slow}}(t, x) = e^{i\omega t} f(x), \quad (1.12)$$

since they lead to breather like (quasi-)periodic solutions for the original equation (1.9), whose average profile, for long  $t \lesssim \varepsilon^{-1}$ , is given by (1.12). Before doing this it turns out to be advantageous to rewrite the non-local non-linear term in (1.11): we define a measure  $\mu(B)$  by setting  $\mu(B) := \frac{1}{2} \int_{-1}^1 \mathbf{1}_B(D(r)) dr$  for any measurable set  $B \subset \mathbb{R}$ . Since  $\mu(B) \geq 0$  and  $\mu(\mathbb{R}) = \int \mathbf{1}_{\mathbb{R}}(\tau) \mu(d\tau) = \int_0^1 \mathbf{1}_{\mathbb{R}}(D(r)) dr = \int_0^1 dr = 1$ , one sees that  $\mu$  is a probability measure. Moreover, as long as  $d_0$  is locally integrable,  $D$  is bounded

and hence the probability measure  $\mu$  also has bounded support. Since  $\mu$  is the image measure of normalized Lebesgue measure on  $[-1, 1]$  under  $D$ , we can rewrite (1.11) as

$$i\partial_t v_{\text{slow}} = d_{\text{av}} A v_{\text{slow}} - \int_{\mathbb{R}} U_{\tau}^{-1} [|U_{\tau} v_{\text{slow}}|^2 U_{\tau} v_{\text{slow}}] \mu(d\tau). \quad (1.13)$$

The simplest case of dispersion management,  $d_0 = 1$  on  $[-1, 0)$  and  $d_0 = -1$  on  $[0, 1]$ , i.e.,  $d_0 = \mathbf{1}_{[-1, 0)} - \mathbf{1}_{[0, 1]}$ , which is the case most studied in the literature, corresponds to the measure  $\mu$  having the density  $\mathbf{1}_{[0, 1]}$ , the uniform distribution on  $[0, 1]$ . More generally, if  $d_0$  is piecewise continuous on  $[-1, 1]$  and bounded away from zero, which is certainly a physically reasonable assumption, or zero on an at most discrete subset of  $[0, 1]$ , the inverse function theorem, [51], shows that  $\mu$  has a density  $\psi$ , i.e.,  $\mu(d\tau) = \psi(\tau) d\tau$ , with  $\psi$  given by

$$\psi(\tau) := \sum_{t \in D^{-1}(\{\tau\})} |d_0(t)|^{-1}, \quad (1.14)$$

where  $D^{-1}(\{\tau\}) = \{t \in [0, 1] \mid D(t) = \tau\}$ .

Finally, we find it convenient to multi-linearize the non-linear and non-local term in (1.13) by introducing

$$Q_{\mu}(v_1, v_2, v_3)(t) := \int_{\mathbb{R}} U_{\tau}^{-1} [U_{\tau} v_1(t, \cdot) \overline{U_{\tau} v_2(t, \cdot)} U_{\tau} v_3(t, \cdot)] \mu(d\tau). \quad (1.15)$$

With this we rewrite (1.11) as

$$i\partial_t v_{\text{slow}} = d_{\text{av}} A v_{\text{slow}} - Q_{\mu}(v_{\text{slow}}, v_{\text{slow}}, v_{\text{slow}}). \quad (1.16)$$

The term  $Q_{\mu}$  is closely related to the non-linear and non-local functionals appearing in the variational problems (1.1) and (1.2). The ansatz (1.12) in (1.16) yields the time independent Gabitov-Turitsyn equation; in our notation,

$$-\omega f = d_{\text{av}} A f - Q_{\mu}(f, f, f), \quad (1.17)$$

which is a non-local non-linear eigenvalue equation for  $f$ . By testing (1.17) with suitable test functions  $g$  one arrives at the weak formulation

$$-\omega \langle g, f \rangle = d_{\text{av}} \langle g, A f \rangle - \langle g, Q_{\mu}(f, f, f) \rangle$$

where  $\langle g, f \rangle$  is either the scalar product on  $L^2(\mathbb{R})$  given by  $\int \overline{g(x)} f(x) dx$  in the continuous case or the scalar product  $\sum_{x \in \mathbb{Z}} \overline{g(x)} f(x)$  on  $l^2(\mathbb{Z})$  in the discrete case. In the continuous case we interpret the quadratic form  $\langle g, A f \rangle$  as  $\langle g, A f \rangle = \langle \partial_x g, \partial_x f \rangle = \langle g', f' \rangle$  by an integration by parts. A formal calculation, using the unicity of  $U_{\tau}$ , yields

$$\langle g, Q_{\mu}(f, f, f) \rangle = \mathcal{Q}_{\mu}^{c/d}(g, f, f, f) \quad (1.18)$$

where the four linear functional  $\mathcal{Q}_{\mu}^{c/d}$  is given by (1.3) in the continuous case and (1.4) in the discrete case, respectively. The formal calculation yielding  $\mathcal{Q}_{\mu}^{c/d}$  is justified in the continuous and discrete case using Lemma B.1, respectively Lemma B.7. Thus the weak formulation of (1.17) is

$$-\omega \langle g, f \rangle = d_{\text{av}} \langle g, A f \rangle - \mathcal{Q}_{\mu}^{c/d}(g, f, f, f), \quad (1.19)$$

supposed to hold for all  $g \in l^2(\mathbb{Z})$  in the discrete case and for any  $g$  in the Sobolev space  $H^1(\mathbb{R})$  in the continuous case.

Equation (1.19) is the weak form of the Euler-Lagrange equation associated with the energy

$$H(f) := \frac{d_{\text{av}}}{2} \langle f, Af \rangle - \frac{1}{4} \mathcal{Q}_\mu^{c/d}(f, f, f, f). \quad (1.20)$$

In particular, any minimizer of the associated constraint minimization problem

$$M_\lambda^{d_{\text{av}}} := \inf(H(f) \mid f \in X, \|f\|_2^2 = \lambda), \quad (1.21)$$

where  $X = l^2(\mathbb{Z})$  in the discrete case and the Sobolev space  $H^1(\mathbb{R})$  in the continuous case, will be, up to some minor technicalities, a solution of (1.19) for some choice of Lagrange multiplier  $\omega$ , as long as the variational problem (1.21) admits minimizers.

The case of vanishing average dispersion/diffraction,  $d_{\text{av}} = 0$ , is of particular practical importance for applications, [15, 16, 56], since in this case the positive and negative dispersion/diffraction exactly cancel out. In the limit  $d_{\text{av}} \rightarrow 0$ , the variational problem (1.21) yields, up to a minor sign change, the two restricted variational problems (1.1), respectively (1.2), in the continuous, respectively discrete, case. Associated with these limiting two variational problems are the Euler-Lagrange equations

$$\omega \langle g, f \rangle = \mathcal{Q}_\mu^c(g, f, f, f), \quad \text{for all } g \in L^2(\mathbb{R}) \quad (1.22)$$

and

$$\omega \langle g, f \rangle = \mathcal{Q}_\mu^d(g, f, f, f), \quad \text{for all } g \in l^2(\mathbb{Z}) \quad (1.23)$$

which, in the language of differential equations, are the *singular limits* of (1.19) for  $d_{\text{av}} = 0$  in the continuous, respectively discrete, case. Note that the continuous version of (1.16) is elliptic for  $d_{\text{av}} > 0$ , whereas its singular limit (1.22) is no longer elliptic. This corresponds to the domain of the variational problem (1.1) increasing from the Sobolev space  $H^1(\mathbb{R})$  to the full space  $L^2(\mathbb{R})$ , due to the loss of second order derivatives.

Solutions of (1.22) are precisely the dispersion management solitons for vanishing average dispersion and solutions of (1.23) the diffraction management solitons for vanishing average diffraction. Theorems 1.1 and 1.2 show that these solitons exist under rather mild conditions on  $\mu$ , translating to very general and non-restrictive conditions on the profile  $d_0$  for existence of dispersion management solitons, see Remark 1.3.i.

As a final remark, we would like to note that even for  $d_{\text{av}} > 0$  existence of minimizers of (1.21) has only been established for the special choice of profile  $d_0 = \mathbf{1}_{[-1,0)} - \mathbf{1}_{[0,1]}$ , corresponding to  $\mu$  having density  $\mathbf{1}_{[0,1]}$  in the continuous case. The continuous case was done in [59] again using Lions' concentration compactness principle. Due to the absence of scaling in  $l^2(\mathbb{Z})$  there is a threshold phenomena for the existence of minimizers of (1.21) in the discrete case similar to [58] which makes the problem slightly harder than the continuous case: The infimum  $M_\lambda^{\text{diff}}$  is negative only for sufficiently large  $\lambda$  and minimizers for (1.21) exist in the discrete case only if  $\lambda$  is large enough, depending on  $d_{\text{av}} > 0$ , see [40] and [44] where the existence of minimizers for piecewise continuous diffraction profile  $d_0$  was shown using a discrete version of Lions' concentration compactness principle.

In all above cases, using by now well-known arguments, see [8, 9, 57], this variational approach to the existence of solutions of (1.19) also shows that minimizers of (1.21)



for  $d_{\text{av}} > 0$ , respectively maximizers of (1.1) and (1.2) for  $d_{\text{av}} = 0$  lead to orbitally stable solutions of (1.16).

## 2. THE EXISTENCE PROOF

We want to show that one can suitably massage any maximizing sequence for the variational problem (1.1), respectively (1.2), with translations and boosts of  $L^2(\mathbb{R})$ , respectively translations of  $l^2(\mathbb{Z})$ , so that it lies in a compact subspace of  $L^2$ , respectively  $l^2$ . The following lemma is the key result for this. First we need one more piece of notation. For  $x > 0$  and  $\alpha \in \mathbb{R}$  define

$$G_\alpha(x) := \left( (x + \alpha^2)^{1/2} - \alpha \right)^{-1/2}. \quad (2.1)$$

Note that  $G_\alpha$  is a decreasing function on  $\mathbb{R}_+$  which vanishes at infinity. Moreover, for  $z \in \mathbb{R}$  let  $z_+ := \max(z, 0)$ .

**Lemma 2.1** (Continuous case). *Let  $\mu$  have a density  $\psi$  satisfying  $0 \leq \psi \in L^2(\mathbb{R}) \cap L^4(\mathbb{R}) \cap L^4(\mathbb{R}, t^2 dt)$ . Then there exists a constant  $C$  depending only on  $\|\psi\|_{L^2(\mathbb{R})}$ ,  $\|\psi\|_{L^4(\mathbb{R})}$  and  $\|\psi\|_{L^4(\mathbb{R}, t^2 dt)}$  such that if  $f \in L^2(\mathbb{R})$ ,  $0 < \varepsilon < \|f\|_{L^2}$ , and  $a, b \in \mathbb{R}$ , with*

$$\int_{-\infty}^a |f(x)|^2 dx \geq \frac{\varepsilon^2}{2} \text{ and } \int_b^\infty |f(x)|^2 dx \geq \frac{\varepsilon^2}{2}, \quad (2.2)$$

then

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq P_1^c(\|f\|_{L^2}^4 - \varepsilon^4/2) + C\|f\|_{L^2}^4 G_{1/2}((b-a)_+). \quad (2.3)$$

Moreover, whenever  $c, d \in \mathbb{R}$ , are such that

$$\int_{-\infty}^c |\widehat{f}(k)|^2 dk \geq \frac{\varepsilon^2}{2} \text{ and } \int_d^\infty |\widehat{f}(k)|^2 dk \geq \frac{\varepsilon^2}{2}, \quad (2.4)$$

then also

$$\mathcal{Q}_\mu^c(\widehat{f}, \widehat{f}, \widehat{f}, \widehat{f}) \leq P_1^c(\|f\|_{L^2}^4 - \varepsilon^4/2) + C\|f\|_{L^2}^4 G_{1/2}((d-c)_+). \quad (2.5)$$

We have a similar bound in the discrete case under much weaker assumptions on the measure  $\mu$ .

**Lemma 2.2** (Discrete case). *Assume that  $\mu$  is a bounded measure with bounded support. Then there is a constant  $C$  such that if  $f \in l^2(\mathbb{Z})$ ,  $0 < \varepsilon < \|f\|_2$  and  $a, b \in \mathbb{Z}$ , with*

$$\sum_{x < a} |f(x)|^2 \geq \frac{\varepsilon^2}{2} \text{ and } \sum_{x > b} |f(x)|^2 \geq \frac{\varepsilon^2}{2}, \quad (2.6)$$

then

$$\mathcal{Q}_\mu^d(f, f, f, f) \leq P_1^d(\|f\|_{l^2}^4 - \varepsilon^4/2) + C\|f\|_{l^2}^4 G_1((b-a+1)_+) \quad (2.7)$$

**Remark 2.3.** Of course, by the monotonicity of  $G_\alpha$ , the bounds (2.3), respectively (2.4), are strongest if one chooses the smallest  $a$ , respectively  $c$ , and the largest  $b$ , respectively  $d$ , under the restrictions (2.2), respectively (2.4). Similarly, the bound (2.7) is strongest if one chooses the smallest  $a$  and the largest  $b$  obeying (2.6). The choice  $a \geq b$  or  $c \geq d$  is allowed, although in this case the bounds do not yield any

information. Unlike the continuum case, where for  $f \in L^2(\mathbb{R})$  and  $0 < \varepsilon < \|f\|_{L^2}$  one can always find  $a < b$  and  $c < d$  such that the bounds (2.2) and (2.4) hold, it can happen in the discrete case that for some  $f \in l^2(\mathbb{Z})$  and  $0 < \varepsilon < \|f\|_{l^2}$  one has  $a \geq b$  or even  $a > b$  for any  $a, b$  obeying (2.6). For example, this happens, even for all  $0 < \varepsilon < \|f\|_{l^2}$ , if  $f \in l^2(\mathbb{Z})$  is concentrated on a single point in  $\mathbb{Z}$ .

Thus the bounds provided by Lemma 2.1 and 2.2 yield the least information for strongly localized functions  $f$ , respectively their Fourier transforms  $\hat{f}$ . At first this might seem counterintuitive since we intend to use these bounds in order to get concentration bounds on  $f$ , respectively  $\hat{f}$ , see Propositions 2.4 and 2.5 below.

*Proof of Lemma 2.1.* We prove only (2.3) since, given Lemma B.3, the proof of (2.5) is nearly identical. We will write  $\|f\|$  for  $\|f\|_{L^2}$  in the following. Of course, since the right hand side of (2.3) is infinite if  $a \geq b$ , we can assume that  $a < b$ . Let  $a'$  and  $b'$  arbitrary numbers with  $a \leq a' < b' \leq b$  and split  $f$  into

$$f = f_{-1} + f_0 + f_1 \quad (2.8)$$

where we set  $f_{-1} = f\mathbf{1}_{(-\infty, a']}$ ,  $f_0 = f\mathbf{1}_{[a', b']}$ , and  $f_1 = f\mathbf{1}_{(b', \infty)}$ . We will choose suitable  $a'$  and  $b'$  soon. Obviously,  $\|f_j\| \leq \|f\|$  for  $j = -1, 0, 1$ . Moreover, since (2.2) and  $a' \geq a$  and  $b' \leq b$  we also have

$$\|f_{-1}\|^2 \text{ and } \|f_1\|^2 \geq \frac{\varepsilon^2}{2}. \quad (2.9)$$

In order to bound  $\mathcal{Q}_\mu^c(f, f, f, f)$  we use its multi-linearity,

$$\begin{aligned} \mathcal{Q}_\mu^c(f, f, f, f) &= \mathcal{Q}_\mu^c(f_{-1} + f_0 + f_1, f, f, f) \\ &= \mathcal{Q}_\mu^c(f_{-1}, f, f, f) + \mathcal{Q}_\mu^c(f_0, f, f, f) + \mathcal{Q}_\mu^c(f_1, f, f, f). \end{aligned} \quad (2.10)$$

The term containing  $f_0$  is simply bounded by

$$|\mathcal{Q}_\mu^c(f_0, f, f, f)| \leq P_1^c \|f_0\| \|f\|^3 \quad (2.11)$$

by Lemma B.1. The other terms we further split into

$$\begin{aligned} \mathcal{Q}_\mu^c(f_1, f, f, f) &= \mathcal{Q}_\mu^c(f_1, f_{-1} + f_0 + f_1, f, f) \\ &= \mathcal{Q}_\mu^c(f_1, f_{-1}, f, f) + \mathcal{Q}_\mu^c(f_1, f_0, f, f) + \mathcal{Q}_\mu^c(f_1, f_1, f, f) \end{aligned} \quad (2.12)$$

with a similar expression for  $\mathcal{Q}_\mu^c(f_{-1}, f, f, f)$ . Again

$$|\mathcal{Q}_\mu^c(f_1, f_0, f, f)| \leq P_1^c \|f_1\| \|f_0\| \|f\|^2 \leq P_1^c \|f_0\| \|f\|^3 \quad (2.13)$$

for the term containing  $f_0$ . Since the supports of  $f_{-1}$  and  $f_1$  have at least distance  $b' - a'$ , the refined multi-linear bound of Lemma B.3 gives estimate

$$|\mathcal{Q}_\mu^c(f_1, f_{-1}, f, f)| \lesssim \frac{\|f_1\| \|f_{-1}\| \|f\|^2}{(b' - a')^{1/2}} \leq \frac{\|f\|^4}{(b' - a')^{1/2}} \quad (2.14)$$

for the first term in (2.12). Terms of the form  $\mathcal{Q}_\mu^c(f_{-1}, f_0, f, f)$  and  $\mathcal{Q}_\mu^c(f_{-1}, f_1, f, f)$  are estimated the same way.

Continuing similarly for the terms  $\mathcal{Q}_\mu^c(f_1, f_1, f, f)$ , respectively  $\mathcal{Q}_\mu^c(f_{-1}, f_{-1}, f, f)$ , we see that there exist a constant  $C < \infty$  such that

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq \mathcal{Q}_\mu^c(f_{-1}, f_{-1}, f_{-1}, f_{-1}) + \mathcal{Q}_\mu^c(f_1, f_1, f_1, f_1)$$

$$\begin{aligned}
& + C \left[ \|f_0\| \|f\|^3 + \frac{\|f\|^4}{(b' - a')^{1/2}} \right] \\
& \leq P_1^c(\|f_{-1}\|^4 + \|f_1\|^4) + C \left[ \|f_0\| \|f\|^3 + \frac{\|f\|^4}{(b' - a')^{1/2}} \right], \tag{2.15}
\end{aligned}$$

where the second inequality follows again from Lemma B.1. Using (2.9), we get

$$\|f_{-1}\|^4 + \|f_1\|^4 = (\|f_{-1}\|^2 + \|f_1\|^2)^2 - 2\|f_{-1}\|^2 \|f_1\|^2 \leq \|f\|^4 - \frac{\varepsilon^4}{2}$$

since  $\|f_{-1}\|^2 + \|f_1\|^2 \leq \|f\|^2$  always. In particular, (2.15) gives

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq P_1^c(\|f\|^4 - \frac{\varepsilon^4}{2}) + C \left[ \|f_0\| \|f\|^3 + \frac{\|f\|^4}{(b' - a')^{1/2}} \right]. \tag{2.16}$$

To choose  $a'$  and  $b'$  let  $0 < l < b - a$  and note that

$$\begin{aligned}
\int_a^{b-l} \int_\eta^{\eta+l} |f(x)|^2 dx d\eta &= \int_{\substack{a \leq \eta \leq b-l \\ \eta \leq x \leq \eta+l}} |f(x)|^2 dx d\eta \leq \int_{\substack{a \leq x \leq b \\ x-l \leq \eta \leq x}} |f(x)|^2 dx d\eta \\
&= l \int_a^b |f(x)|^2 dx \leq l \|f\|^2. \tag{2.17}
\end{aligned}$$

By the mean value theorem and (2.17), there exists  $\eta' \in (a, b-l)$  such that

$$(b - a - l) \int_{\eta'}^{\eta'+l} |f(x)|^2 dx = \int_a^{b-l} \int_\eta^{\eta+l} |f(x)|^2 dx d\eta.$$

Thus, with the choice  $a' = \eta'$  and  $b' = \eta' + l$  we have  $b' - a' = l$ ,  $a < a' < b' < b$ , and

$$\|f_0\|^2 \leq \frac{l}{b - a - l} \|f\|^2.$$

Plugging this into (2.16) yields

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq P_1^c(\|f\|^4 - \frac{\varepsilon^4}{2}) + C \|f\|^4 \left[ \left( \frac{l}{b - a - l} \right)^{1/2} + \frac{1}{l^{1/2}} \right] \tag{2.18}$$

for any  $0 < \varepsilon < \|f\|$  and all  $0 < l < b - a$ . The choice  $l = \sqrt{b - a + 1/4} - 1/2$ , which is allowed since  $0 < \sqrt{s + 1/4} - 1/2 < s$  for any  $s > 0$ , gives

$$\frac{l}{b - a - l} = \frac{1}{l}.$$

Hence (2.18) yields (2.3). ■

*Proof of Lemma 2.2.* As in the continuous case, we can assume that  $a, b \in \mathbb{Z}$  with  $a \leq b$  and we write  $\|f\|$  for the  $l^2$ -norm of a function  $f \in l^2(\mathbb{Z})$ . Again, for any choice  $a \leq a' \leq b' \leq b$  one puts  $f_{-1} = f \mathbf{1}_{(-\infty, a')} = f \mathbf{1}_{(-\infty, a'-1]}$ ,  $f_1 = f \mathbf{1}_{(b', \infty)} = f \mathbf{1}_{[b'+1, \infty)}$ , and  $f_0 = f \mathbf{1}_{[a', b']}$ . Furthermore, let  $l$  be the number of points in  $[a', b']$ , i.e.,  $l = b' - a' + 1$  and note that  $\text{dist}(\text{supp}(f_{-1}), \text{supp}(f_1)) = l + 1$ . Then we argue exactly as in the continuous case, but use the bounds from Lemma B.7 and B.8 instead, to see that

$$\mathcal{Q}_\mu^d(f, f, f, f) \leq P_1^d(\|f\|_2^4 - \frac{\varepsilon^4}{2}) + C \left[ \|f_0\|_2 \|f\|_2^3 + \|f\|_2^4 (l + 1)^{-(l+1)/4} \right] \tag{2.19}$$

holds. The discrete version of 2.17 now reads

$$\sum_{\eta=a}^{b-l+1} \sum_{x=\eta}^{\eta+l-1} |f(x)|^2 \leq \sum_{x=a}^b \sum_{\eta=x-l+1}^x |f(x)|^2 \leq l \|f\|_2^2. \quad (2.20)$$

By pidgeonholing, since the number of points in  $[a, b-l+1]$  is  $b-a+2-l$ , there must exists  $\eta'$  with  $a \leq \eta' \leq b-l+1$  and

$$(b-a+2-l) \sum_{x=\eta'}^{\eta'+l-1} |f(x)|^2 \leq \sum_{\eta=a}^{b-l+1} \sum_{x=\eta}^{\eta+l-1} |f(x)|^2.$$

Thus, choosing  $a' = \eta'$  and  $b' = \eta' + l - 1$ , the bounds (2.19) and (2.20) give

$$\mathcal{Q}_\mu^d(f, f, f, f) \leq P_1^d(\|f\|_2^4 - \frac{\varepsilon^4}{2}) + C \|f\|_2^4 \left[ \left( \frac{l}{b-a+2-l} \right)^{1/2} + (l+1)^{-(l+1)/4} \right] \quad (2.21)$$

Now there exists  $l \in \mathbb{N}$  with  $l \leq (b-a+2)^{1/2} < l+1$ , which is an allowed choice for  $l$ , i.e., it obeys  $1 \leq l \leq b-a+1$  for any  $a \leq b \in \mathbb{Z}$ . With this choice the estimates

$$\left( \frac{l}{b-a+2-l} \right)^{1/2} \leq \left( \frac{(b-a+2)^{1/2}}{b-a+2-(b-a+2)^{1/2}} \right)^{1/2} = ((b-a+2)^{1/2} - 1)^{-1/2}$$

and, since  $l \geq 1$ ,

$$(l+1)^{-(l+1)/4} \leq (l+1)^{-1/2} \leq (b-a+2)^{-1/4} \leq ((b-a+2)^{1/2} - 1)^{-1/2}$$

show that (2.21) implies (2.7). ■

Lemma 2.1 has strong consequences for maximizing sequences of the variational problem (1.1). Recall that  $(f_n)_n \subset L^2(\mathbb{R})$  is a maximizing sequence for (1.1) if  $\|f_n\|_{L^2}^2 = \lambda > 0$  for all  $n \in \mathbb{N}$  and

$$\lim_{n \rightarrow \infty} \mathcal{Q}_\mu^c(f_n, f_n, f_n, f_n) = \sup \left( \mathcal{Q}_\mu^c(g, g, g, g) \mid g \in L^2(\mathbb{R}), \|g\|_{L^2}^2 = \lambda \right).$$

The following proposition shows that any maximizing sequence for the variational problem (1.1) is *tight modulo translations and shifts*.

**Proposition 2.4** (Tightness, continuous case). *Let  $(f_n)_{n \in \mathbb{N}} \subset L^2(\mathbb{R})$  be a maximizing sequence for the variational problem (1.1) with  $\lambda = \|f_n\|_2^2 > 0$ . Then there exist shifts  $\xi_n$  and boosts  $v_n$  such that*

$$\lim_{R \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{|x-\xi_n| > R} |f_n(x)|^2 dx = 0, \quad (2.22)$$

and

$$\lim_{L \rightarrow \infty} \sup_{n \rightarrow \infty} \int_{|k-v_n| > L} |\widehat{f_n}(k)|^2 dk = 0. \quad (2.23)$$

*Proof.* We will only prove (2.22) since, given the bound (2.5) in Lemma 2.1, the proof of (2.23) is identical. To prove (2.22), we have to show that there exist shifts  $\xi_n$  such that for any  $0 < \varepsilon < \sqrt{\lambda}$  there exists  $R_\varepsilon < \infty$  with

$$\sup_{n \in \mathbb{N}} \int_{|x-\xi_n| > R_\varepsilon} |f_n(x)|^2 dx \leq \varepsilon^2. \quad (2.24)$$

Define

$$a_{n,\varepsilon} := \inf \left( a \in \mathbb{R} : \int_{-\infty}^a |f_n(x)|^2 dx \geq \frac{\varepsilon^2}{2} \right) - 1 \quad (2.25)$$

and

$$b_{n,\varepsilon} := \sup \left( b \in \mathbb{R} : \int_b^\infty |f_n(x)|^2 dx \geq \frac{\varepsilon^2}{2} \right) + 1. \quad (2.26)$$

Both  $a_{n,\varepsilon}$  and  $b_{n,\varepsilon}$  exist and are finite since  $f_n \in L^2(\mathbb{R})$  and  $f_n \not\equiv 0$ . Moreover,  $a_{n,\varepsilon} < b_{n,\varepsilon}$  for all  $n \in \mathbb{N}$  and  $0 < \varepsilon < \sqrt{\lambda}$ , and they are monotone:  $a_{n,\varepsilon_1} \leq a_{n,\varepsilon_2}$  and  $b_{n,\varepsilon_1} \geq b_{n,\varepsilon_2}$  for all  $0 < \varepsilon_1 \leq \varepsilon_2 < \sqrt{\lambda}$  and  $n \in \mathbb{N}$ . Fix  $0 < \varepsilon_0 < \sqrt{\lambda}$  and put

$$\xi_n := (b_{n,\varepsilon_0} + a_{n,\varepsilon_0})/2, \quad (2.27)$$

or choose any point in  $[a_{n,\varepsilon_0}, b_{n,\varepsilon_0}]$ , and assume, for now, that

$$R_\varepsilon := \sup_{n \in \mathbb{N}} (b_{n,\varepsilon} - a_{n,\varepsilon}) < \infty \quad (2.28)$$

for  $0 < \varepsilon \leq \varepsilon_0$  and put  $R_\varepsilon = R_{\varepsilon_0}$  for  $\varepsilon_0 < \varepsilon < \sqrt{\lambda}$ . With this and (2.25) and (2.26) one can easily check that

$$\int_{|x-\xi_n|>R_\varepsilon} |f_n(x)|^2 dx \leq \int_{-\infty}^{a_{n,\varepsilon}} |f_n(x)|^2 dx + \int_{b_{n,\varepsilon}}^\infty |f_n(x)|^2 dx \leq \varepsilon^2$$

for all  $n \in \mathbb{N}$ , which proves (2.24) if we can show that  $R_\varepsilon$  defined in (2.28) is indeed finite. This is where the bound (2.3) of Lemma 2.1 enters.

By our choice of  $a_{n,\varepsilon}$  and  $b_{n,\varepsilon}$  in (2.25) and (2.26) we have

$$\int_{-\infty}^{a_{n,\varepsilon}+2} |f_n(x)|^2 dx \geq \frac{\varepsilon^2}{2} \quad \text{and} \quad \int_{b_{n,\varepsilon}-2}^\infty |f_n(x)|^2 dx \geq \frac{\varepsilon^2}{2}.$$

Thus with  $a = a_{n,\varepsilon} + 2$  and  $b = b_{n,\varepsilon} - 2$  the assumption (2.2) of Lemma 2.1 is fulfilled. Putting  $R_{n,\varepsilon} = b_{n,\varepsilon} - a_{n,\varepsilon}$ , using the scaling  $P_1^c \|f_n\|_2^4 = P_1^c \lambda^2 = P_\lambda$ , see (B.1) in the appendix, and rearranging (2.3) a bit, yields

$$P_1^c \frac{\varepsilon^4}{2} + \mathcal{Q}_\mu^c(f_n, f_n, f_n, f_n) - P_\lambda \leq C\lambda^2 G_{1/2}((R_{n,\varepsilon} - 4)_+). \quad (2.29)$$

Since  $f_n$  is a maximizing sequence for  $P_\lambda$  we have  $\lim_{n \rightarrow \infty} \mathcal{Q}_\mu^c(f_n, f_n, f_n, f_n) = P_\lambda$ . The function  $G_{1/2}$  is decreasing on  $\mathbb{R}_+$ . Thus

$$\liminf_{n \rightarrow \infty} G_{1/2}((R_{n,\varepsilon} - 4)_+) = G_{1/2}((\limsup_{n \rightarrow \infty} R_{n,\varepsilon} - 4)_+).$$

Hence taking the limit  $n \rightarrow \infty$  in (2.29) one sees

$$0 < P_1^c \frac{\varepsilon^4}{2} \leq C\lambda^2 \liminf_{n \rightarrow \infty} G_{1/2}((R_{n,\varepsilon} - 4)_+) = C\lambda^2 G_{1/2}((\limsup_{n \rightarrow \infty} R_{n,\varepsilon} - 4)_+). \quad (2.30)$$

Since  $G_{1/2}$  goes to zero at infinity, the bound (2.30) shows that

$$\limsup_{n \rightarrow \infty} R_{n,\varepsilon} < \infty \quad \text{for all } 0 < \varepsilon < \sqrt{\lambda} \quad (2.31)$$

which proves (2.28) and hence the Lemma. ■

Of course, we have an analogous proposition for any maximizing sequence of the discrete variational problem (1.2).

**Proposition 2.5** (Discrete case). *Let  $(f_n)_{n \in \mathbb{N}} \subset l^2(\mathbb{Z})$  be a maximizing sequence for the variational problem (1.2) with  $\lambda = \|f_n\|_2^2$ . Then there exist shifts  $\xi_n \in \mathbb{Z}$*

$$\lim_{R \rightarrow \infty} \sup_{n \in \mathbb{N}} \sum_{|x - \xi_n| > R} |f_n(x)|^2 = 0 \quad (2.32)$$

*Proof.* Given Lemma 2.2, the proof of Proposition 2.5 is virtually identical to the proof of Proposition 2.4.  $\blacksquare$

Propositions 2.4 and 2.5 are key to our proof of the existence of maximizers.

*Proof of Theorem 1.1 and 1.2:* We only give the details in the continuous case. The discrete case follows by a similar reasoning.

The idea is to use Proposition 2.4 and Lemma A.1 in order to massage an arbitrary maximizing sequence into a strongly convergent sequence. Its strong limit will then furnish the sought after maximizer.

Let  $(f_n)_n \subset L^2(\mathbb{R})$  be an arbitrary maximizing sequence of the variational problem (1.1). Proposition 2.4 guarantees the existence of shifts  $\xi_n \in \mathbb{R}$  and boosts  $v_n \in \mathbb{R}$  such that (2.22) and (2.23) hold. Define the shifted and boosted sequence

$$\tilde{f}_n(x) := (e^{ixv_n} e^{-i\xi_n P} f_n)(x) = e^{iv_n} f_n(x - \xi_n).$$

where  $P = -i\partial_x$  is the one-dimensional momentum operator.

Note that  $\|\tilde{f}_n\|_2^2 = \|f_n\|_2^2 = \lambda$  since shifts and boost are unitary operations on  $L^2(\mathbb{R})$ . As discussed in the introduction, due to the Galilei covariance of the free Schrödinger equation, see (C.5), the functional  $\mathcal{Q}_\mu^c$  is invariant under shifts and boosts, i.e.,  $\mathcal{Q}_\mu^c(\tilde{f}_n, \tilde{f}_n, \tilde{f}_n, \tilde{f}_n) = \mathcal{Q}_\mu^c(f, f, f, f)$ . Hence  $(\tilde{f}_n)_n$  is also a maximizing sequence.

Certainly  $|\tilde{f}_n(x)| = |f_n(x - \xi_n)|$  for all  $n \in \mathbb{N}$ . The Fourier transform of  $\tilde{f}_n$  is given by

$$\widehat{\tilde{f}_n}(k) = \frac{1}{\sqrt{2\pi}} \int e^{-ikx} e^{ixv_n} f_n(x - \xi_n) dx = e^{-i\xi_n(k - v_n)} \widehat{f}_n(k - v_n). \quad (2.33)$$

Thus also  $|\widehat{\tilde{f}_n}(k)| = |\widehat{f}_n(k - v_n)|$ . In particular, (2.22) and (2.23) show that the maximizing sequence  $(\tilde{f}_n)_n$  is tight in the sense of Lemma A.1.

Since  $(\tilde{f}_n)_n$  is bounded in  $L^2(\mathbb{R})$ , the weak compactness of the unit ball, [35], guarantees the existence of a weakly converging subsequence  $(\tilde{f}_{n_j})_j$  of  $(\tilde{f}_n)_n$ . Obviously, this subsequence is also tight in the sense of Lemma A.1 and hence converges even strongly in  $L^2$ . We set

$$f = \lim_{j \rightarrow \infty} \tilde{f}_{n_j}.$$

By strong convergence  $\|f\|_2^2 = \lim_{j \rightarrow \infty} \|\tilde{f}_{n_j}\|_2^2 = \lambda$ . To conclude that  $f$  is the sought after maximizer we note that by the following Lemma 2.6 the map  $f \mapsto \mathcal{Q}_\mu^c(f, f, f, f)$  is continuous on  $L^2(\mathbb{R})$ . Hence

$$\mathcal{Q}_\mu^c(f, f, f, f) = \lim_{j \rightarrow \infty} \mathcal{Q}_\mu^c(\tilde{f}_{n_j}, \tilde{f}_{n_j}, \tilde{f}_{n_j}, \tilde{f}_{n_j}) = P_\lambda.$$

where the last equality follows since  $(\tilde{f}_n)_n$  is a maximizing sequence. Thus  $f$  is a maximizer for the variational problem (1.1).

The proof that the above maximizer is a weak solution of the associated Euler–Lagrange equation (1.22) is standard in the calculus of variations, we sketch it for the convenience of the reader: Let  $\varphi(f) = \varphi_{\text{cont}}(f) = \mathcal{Q}_\mu^c(f, f, f, f)$ . Lemma 2.7 below shows that the derivative of the functional  $\varphi$  at any  $f \in L^2(\mathbb{R})$  is given by the linear map  $D\varphi(f)[h] = 4\text{Re}\mathcal{Q}_\mu^c(h, f, f, f)$ . Similarly, one can check that the derivative of  $\psi(f) = \|f\|_2^2 = \langle f, f \rangle$  is given by  $D\psi(f)[h] = 2\text{Re}\langle h, f \rangle$ . Note that although, in our convention for the inner product, the map  $h \mapsto \langle h, f \rangle$  is anti-linear, the map  $h \mapsto \text{Re}\langle h, f \rangle$  is linear. Similarly, one easily checks that for fixed  $f$  the map  $h \mapsto \text{Re}\mathcal{Q}_\mu^c(h, f, f, f)$  is linear.

Now let  $f$  be any maximizer of the constraint variational problem (1.1) and  $h \in L^2(\mathbb{R})$  arbitrary. Define, for any  $(s, t) \in \mathbb{R}^2$ ,

$$\begin{aligned} F(s, t) &:= \varphi(f + sf + th), \\ G(s, t) &:= \psi(f + sf + th). \end{aligned}$$

Note that

$$\begin{aligned} \nabla F(s, t) &= \begin{pmatrix} D\varphi(f + sf + th)[f] \\ D\varphi(f + sf + th)[h] \end{pmatrix} \\ &= 4 \begin{pmatrix} \text{Re}\mathcal{Q}_\mu^c(f, f + sf + th, f + sf + th, f + sf + th) \\ \text{Re}\mathcal{Q}_\mu^c(h, f + sf + th, f + sf + th, f + sf + th) \end{pmatrix} \end{aligned}$$

and

$$\nabla G(s, t) = \begin{pmatrix} D\psi(f + sf + th)[f] \\ D\psi(f + sf + th)[h] \end{pmatrix} = 2 \begin{pmatrix} \text{Re}\langle f, f + sf + th \rangle \\ \text{Re}\langle h, f + sf + th \rangle \end{pmatrix}.$$

Since  $\langle f, f \rangle = \lambda \neq 0$ ,

$$\nabla G(0, 0) = 2 \begin{pmatrix} \langle f, f \rangle \\ \text{Re}\langle h, f \rangle \end{pmatrix}$$

is not the zero vector in  $\mathbb{R}^2$  and since  $\nabla G(s, t)$  depends multi-linearly, in particular continuously, on  $(s, t)$ , the implicit function theorem [51] shows that there exists an open interval  $I \subset \mathbb{R}$  containing 0 and a differentiable function  $\phi$  on  $I$  with  $\phi(0) = 0$  such that

$$\lambda = \|f\|_2^2 = G(0, 0) = G(\phi(t), t)$$

for all  $t \in I$ . Consider the function  $I \ni t \mapsto F(\phi(t), t)$ . Since  $f$  is a maximizer for the constraint variational problem (1.1),  $F(\phi(t), t)$  has a local maximum at  $t = 0$ . Hence, using the chain rule,

$$0 = \left. \frac{dF(\phi(t), t)}{dt} \right|_{t=0} = \nabla F(0, 0) \cdot \begin{pmatrix} \phi'(0) \\ 1 \end{pmatrix} = 4\mathcal{Q}_\mu^c(f, f, f, f)\phi'(0) + 4\text{Re}\mathcal{Q}_\mu^c(h, f, f, f).$$

Since  $\lambda = G(\phi(t), t)$ , the chain rule also yields

$$0 = \left. \frac{dG(\phi(t), t)}{dt} \right|_{t=0} = \nabla G(0, 0) \cdot \begin{pmatrix} \phi'(0) \\ 1 \end{pmatrix} = 2\langle f, f \rangle\phi'(0) + 2\text{Re}\langle h, f \rangle.$$

Solving this for  $\phi'(0)$  and plugging it back into the expression for the derivative of  $F$ , we see that

$$\frac{\mathcal{Q}_\mu^c(f, f, f, f)}{\langle f, f \rangle} \operatorname{Re} \langle h, f \rangle = \operatorname{Re} \mathcal{Q}_\mu^c(h, f, f, f).$$

In other words, with  $\omega := \mathcal{Q}_\mu^c(f, f, f, f)/\langle f, f \rangle = P_\lambda/\lambda > 0$  and  $f$  any maximizer of (1.1), we have

$$\operatorname{Re}(\omega \langle h, f \rangle) = \operatorname{Re} \mathcal{Q}_\mu^c(h, f, f, f) \quad (2.34)$$

for any  $h \in L^2(\mathbb{R})$ . Replacing  $h$  by  $ih$  in (2.34), one gets

$$\operatorname{Im}(\omega \langle h, f \rangle) = \operatorname{Im} \mathcal{Q}_\mu^c(h, f, f, f) \quad (2.35)$$

for all  $h \in L^2(\mathbb{R})$ . (2.34) and (2.35) together show

$$\omega \langle h, f \rangle = \mathcal{Q}_\mu^c(h, f, f, f)$$

for any  $h \in L^2(\mathbb{R})$ , that is,  $f$  is a weak solution of the dispersion management equation (1.22).  $\blacksquare$

**Lemma 2.6.** (i) *If the measure  $\mu$  has density  $\psi \in L^2(\mathbb{R})$  then the map  $L^2(\mathbb{R}) \ni f \mapsto \mathcal{Q}_\mu^c(f, f, f, f)$  is locally Lipschitz continuous on  $L^2(\mathbb{R})$ .*

(ii) *If the measure  $\mu$  is bounded then the map  $l^2(\mathbb{Z}) \ni f \mapsto \mathcal{Q}_\mu^d(f, f, f, f)$  is locally Lipschitz continuous on  $l^2(\mathbb{Z})$ .*

*Proof.* Using the multi-linearity of  $\mathcal{Q}_\mu^c$ , given  $f, g \in L^2(\mathbb{R})$ , one has

$$\begin{aligned} \varphi(f) - \varphi(g) &= \mathcal{Q}_\mu^c(f, f, f, f) - \mathcal{Q}_\mu^c(g, g, g, g) \\ &= \mathcal{Q}_\mu^c(f - g, f, f, f) + \mathcal{Q}_\mu^c(g, f - g, f, f) + \mathcal{Q}_\mu^c(g, g, f - g, f) + \mathcal{Q}_\mu^c(g, g, g, f - g) \end{aligned} \quad (2.36)$$

This together with the triangle inequality and the a-priori bound of Lemma B.1 immediately yields

$$\begin{aligned} |\varphi(f) - \varphi(g)| &\leq P_1^c \sum_{j=0}^3 \|f\|_2^{3-j} \|f - g\|_2 \|g\|_2^j \\ &\leq 4P_1^c \max(1, \|f\|_2^3, \|g\|_2^3) \|f - g\|_2. \end{aligned} \quad (2.37)$$

The discrete case is proven the same way using Lemma B.7  $\blacksquare$

In the proof of Theorems 1.1 and 1.2 we needed one more technical result, about the differentiability of the non-linear functionals  $\mathcal{Q}_\mu^{c/d}$ :

**Lemma 2.7.** (i) *If the measure  $\mu$  has density  $\psi \in L^2(\mathbb{R})$  then the map  $L^2(\mathbb{R}) \ni f \mapsto \varphi_\mu^c(f) = \mathcal{Q}_\mu^c(f, f, f, f)$  is continuously differentiable with derivative  $D\varphi_\mu^c(f)[h] = 4\operatorname{Re} \mathcal{Q}_\mu^c(h, f, f, f)$ .*

(ii) *If the measure  $\mu$  is bounded then the map  $l^2(\mathbb{Z}) \ni f \mapsto \varphi_\mu^d(f) = \mathcal{Q}_\mu^d(f, f, f, f)$  is continuously differentiable with derivative  $D\varphi_\mu^d(f)[h] = 4\operatorname{Re} \mathcal{Q}_\mu^d(h, f, f, f)$ .*



*Proof.* Using the multi-linearity of  $\mathcal{Q}_\mu^c$ , one can check that for any  $f, h \in L^2(\mathbb{R})$

$$\begin{aligned} \varphi_\mu^c(f+h) &= \varphi_\mu^c(f) + \mathcal{Q}_\mu^c(f, f, f, h) + \mathcal{Q}_\mu^c(f, f, h, f) + \mathcal{Q}_\mu^c(f, h, f, f) + \mathcal{Q}_\mu^c(h, f, f, f) + O(\|h\|_2^2) \\ &= \varphi_\mu^c(f) + 4\operatorname{Re}\mathcal{Q}_\mu^c(h, f, f, f) + O(\|h\|_2^2) \end{aligned} \quad (2.38)$$

where in the term  $O(\|h\|_2^2)$  we gathered expressions of the form  $\mathcal{Q}_\mu^c(h, f, f, h)$ , and  $\mathcal{Q}_\mu^c(h, h, f, f)$  or  $\mathcal{Q}_\mu^c(h, h, f, f+h)$  and similar which, by the a-priori bound from Lemma B.1, are bounded by  $C\|h\|_2^2$  with  $C \leq P_1^c \max(1, \|f\|_2^2, \|h\|_2^2)$ . This shows that  $\varphi_\mu^c$  is differentiable with derivative  $D\varphi_\mu^c(f)[h] = 4\operatorname{Re}\mathcal{Q}_\mu^c(h, f, f, f)$ . Moreover,

$$\begin{aligned} D\varphi_\mu^c(f)[h] - D\varphi_\mu^c(g)[h] &= 4\operatorname{Re}(\mathcal{Q}_\mu^c(h, f, f, f) - \mathcal{Q}_\mu^c(h, g, g, g)) \\ &= 4\operatorname{Re}(\mathcal{Q}_\mu^c(h, f - g, f, f) + \mathcal{Q}_\mu^c(h, g, f - g, f) + \mathcal{Q}_\mu^c(h, g, g, f - g)). \end{aligned}$$

Hence using the bound from Lemma B.1 again, we see

$$\sup_{\|h\|_2 \leq 1} |D\varphi_\mu^c(f)[h] - D\varphi_\mu^c(g)[h]| \lesssim (\|f\|_2^2 + \|f\|_2\|g\|_2 + \|g\|_2^2)\|f - g\|_2 \quad (2.39)$$

which shows that the derivative  $D\varphi_\mu^c$  is even locally Lipschitz continuous. The discrete case is proven analogously.  $\blacksquare$

## APPENDIX A. STRONG CONVERGENCE

A key step in our existence proof of maximizers of the variational problems (1.1) and (1.2) is the characterization of strong convergence in  $L^2(\mathbb{R})$  and  $l^2(\mathbb{Z})$ . We need only the respective one-dimensional versions, but give the result and its proof for all dimensions. In the discrete case the simple characterization of strong convergence extends to  $l^p(\mathbb{Z}^d)$  for any  $1 \leq p < \infty$ . We start with

**Lemma A.1.** *A sequence  $(f_n)_{n \in \mathbb{N}} \subset L^2(\mathbb{R}^d)$  is strongly converging to  $f$  in  $L^2(\mathbb{R}^d)$  if and only if it is weakly convergent to  $f$  and*

$$\lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_{|x| > R} |f_n(x)|^2 dx = 0, \quad (A.1)$$

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_{|k| > L} |\widehat{f}_n(k)|^2 dk = 0, \quad (A.2)$$

where  $\widehat{f}$  is the Fourier transform of  $f$ .

**Remarks A.2.** (i) A bounded sequence in  $L^2(\mathbb{R}^d)$  can converge weakly to zero by vanishing, splitting, or oscillating to death. (A.1) prevents splitting and vanishing and (A.2) prevents oscillating to death, which corresponds to vanishing or splitting in Fourier space.

(ii) Given a sequence  $(f_n)_{n \in \mathbb{N}} \subset L^2(\mathbb{R}^d)$  the bounds (A.1) and (A.2) certainly hold if there exists functions  $H, F \geq 1$  with  $\lim_{|x| \rightarrow \infty} H(x) = \infty = \lim_{|k| \rightarrow \infty} F(k)$  and

$$\limsup_{n \rightarrow \infty} \int H(x) |f_n(x)|^2 dx < \infty, \quad (A.3)$$

$$\limsup_{n \rightarrow \infty} \int F(k) |\widehat{f_n}(k)|^2 dk < \infty, \quad (\text{A.4})$$

on the other hand, it is easy to see that once (A.1) holds then there exist a function  $H$  bounded below by one and growing to infinity at infinity such that (A.3) holds. So (A.1) and (A.3), and hence also (A.2) and (A.4) are equivalent. In particular, in the proof of the difficult part of Lemma A.1 one could use Rellich's compactness result, see [45]. We prefer, however, the proof given below, which uses only simple properties of compact operators.

(iii) Of course Lemma A.1 holds also with  $\limsup_{n \rightarrow \infty}$  replaced by  $\sup_{n \in \mathbb{N}}$ .

*Proof.* Assume that  $f_n$  converges to  $f$  strongly. Then it certainly converges weakly to  $f$ , that is,  $\lim_{n \rightarrow \infty} \langle \varphi, f_n \rangle = \langle \varphi, f \rangle$  for all  $\varphi \in L^2(\mathbb{R}^d)$ . To check the two conditions (A.1) and (A.2), let  $B_R(0) = \{x \in \mathbb{R}^d : |x| \leq 1\}$  be the closed ball of radius  $R$  and for a self-adjoint (vector-)operator  $A$  let  $\chi_R(A) = \mathbf{1}_{B_R(0)}(A)$  be the associated orthogonal spectral projection and  $\overline{\chi_R}(A) := \mathbf{1} - \chi_R(A)$  the orthogonal projection onto the orthogonal complement of  $\text{ran}(\chi_R(A))$ . Then, for any  $g \in L^2(\mathbb{R}^d)$ ,

$$\int_{|x| > R} |g(x)|^2 dx = \|\overline{\chi_R}(X)g\|_2^2$$

where  $X$  is the position operator, i.e., multiplication by  $x$ . Since  $\overline{\chi_R}(X)$  is an orthogonal projection in  $L^2(\mathbb{R}^d)$ , the triangle inequality yields

$$\|\overline{\chi_R}(X)f_n\| \leq \|\overline{\chi_R}(X)f\| + \|\overline{\chi_R}(X)(f_n - f)\| \leq \|\overline{\chi_R}(X)f\| + \|f_n - f\|$$

So, since  $f_n$  converges in norm to  $f$ ,

$$\limsup_{n \rightarrow \infty} \|\overline{\chi_R}(X)f_n\| \leq \|\overline{\chi_R}(X)f\|$$

for all  $R > 0$  and (A.1) follows by taking the limit  $R \rightarrow \infty$ , since  $\overline{\chi_R}(X)$  converges strongly to zero as  $R \rightarrow \infty$ . (A.2) follows by an identical argument, using

$$\int_{|k| > L} |\widehat{g}(k)|^2 dk = \|\overline{\chi_L}(P)g\|^2$$

with  $P = -i\nabla$  the momentum operator in  $L^2(\mathbb{R}^d)$ . As above, we see

$$\limsup_{n \rightarrow \infty} \|\overline{\chi_L}(P)f_n\| \leq \|\overline{\chi_L}(P)f\|$$

which implies (A.2) in the limit  $L \rightarrow \infty$ .

For the converse assume that  $f_n$  converges weakly to  $f \in L^2(\mathbb{R}^d)$  and that (A.1) and (A.2) hold. A judicious use of the triangle inequality reveals

$$\begin{aligned} \|f - f_n\| &\leq \|\chi_R(X)(f - f_n)\| + \|\overline{\chi_R}(X)(f - f_n)\| \\ &\leq \|\chi_R(X)\chi_L(P)(f - f_n)\| + \|\chi_R(X)\overline{\chi_L}(P)(f - f_n)\| + \|\overline{\chi_R}(X)(f - f_n)\| \\ &\leq \|\chi_R(X)\chi_L(P)(f - f_n)\| + \|\overline{\chi_L}(P)(f - f_n)\| + \|\overline{\chi_R}(X)(f - f_n)\| \\ &\leq \|\chi_R(X)\chi_L(P)(f - f_n)\| + \|\overline{\chi_L}(P)f\| + \|\overline{\chi_R}(X)f\| + \|\overline{\chi_L}(P)f_n\| + \|\overline{\chi_R}(X)f_n\| \end{aligned}$$

Note that  $\chi_R(X)\chi_L(P)$  is a Hilbert-Schmidt operator, in particular, a compact operator, see, for example, [5, 47]. Thus it maps weakly convergent sequences into strongly

convergent sequences. Hence  $\lim_{n \rightarrow \infty} \|\chi_R(X)\chi_L(P)(f - f_n)\| = 0$  since  $f_n$  converges weakly to  $f$ . Taking the limit  $n \rightarrow \infty$  in the above inequality one sees

$$\limsup_{n \rightarrow \infty} \|f - f_n\| \leq \|\overline{\chi_L}(P)f\| + \|\overline{\chi_R}(X)f\| + \limsup_{n \rightarrow \infty} \|\overline{\chi_L}(P)f_n\| + \limsup_{n \rightarrow \infty} \|\overline{\chi_R}(X)f_n\|$$

and then taking the limit  $L, R \rightarrow \infty$  using  $f \in L^2(\mathbb{R}^d)$  together with (A.1) and (A.2), we get

$$\limsup_{n \rightarrow \infty} \|f - f_n\| = 0,$$

that is,  $f_n$  converges strongly to  $f$ . ■

**Remark A.3.** That  $\chi_R(X)\chi_L(P)$  is Hilbert-Schmidt is easy to see: Using the Fourier transform, one sees that  $\chi_R(X)\chi_L(P)$  has an integral kernel with

$$\chi_R(X)\chi_L(P)(x, y) = \frac{1}{(2\pi)^{d/2}} \chi_R(x) \widehat{\chi_L}(x - y). \quad (\text{A.5})$$

Thus the Hilbert-Schmidt norm of  $\chi_R(X)\chi_L(P)$  is given by

$$\|\chi_R(X)\chi_L(P)\|_{\text{HS}}^2 := \iint_{\mathbb{R}^d \times \mathbb{R}^d} |\chi_R(X)\chi_L(P)(x, y)|^2 dx dy = (2\pi)^{-d/4} \|\chi_R\|_{L^2}^2 \|\chi_L\|_{L^2}^2$$

which is finite for all  $0 < R, L < \infty$ .

The discrete version of the strong convergence result is

**Lemma A.4.** *Let  $1 \leq p < \infty$ . A sequence  $(f_n)_{n \in \mathbb{N}} \subset l^p(\mathbb{Z}^d)$  is strongly converging to  $f$  in  $l^p(\mathbb{Z}^d)$  if and only if it is weakly convergent to  $f$  and the sequence is tight, i.e.,*

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{|x| > L} |f_n(x)|^p = 0. \quad (\text{A.6})$$

*Proof.* The proof uses the same ideas in the continuous case. Let  $K_L$  denote the operator of multiplication with the characteristic function of the set  $\{x \in \mathbb{Z}^d : |x| \leq L\}$ , that is,

$$K_L f(x) = \begin{cases} f(x) & \text{for } |x| \leq L \\ 0 & \text{for } |x| > L \end{cases}$$

and  $\overline{K}_L := \mathbf{1} - K_L$ . Note that for all  $L \geq 1$  both  $K_L$  and  $\overline{K}_L$  are bounded operators on  $l^p(\mathbb{Z}^d)$  with operator norm one since

$$\|f\|_p^p = \|K_L f\|_p^p + \|\overline{K}_L f\|_p^p \geq [\max(\|K_L f\|_p, \|\overline{K}_L f\|_p)]^p$$

for all  $f \in l^p(\mathbb{Z}^d)$ . Moreover, the tightness-condition (A.6) is equivalent to

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\overline{K}_L f_n\|_p = 0. \quad (\text{A.7})$$

Since, for  $1 \leq p < \infty$ ,  $\overline{K}_L$  converges strongly to zero as  $L \rightarrow \infty$ , the proof of Lemma A.4 is then as for the continuous case. In fact, it is much simpler since  $K_L$  is even a finite range operator and thus trivially maps weakly converging sequences into strongly convergent sequences. ■

**Remark A.5.** The proof above breaks down for  $p = \infty$ , since for an arbitrary  $f \in l^\infty(\mathbb{Z}^d)$ , one does not have

$$\lim_{L \rightarrow \infty} \|\overline{K}_L f\|_\infty = 0,$$

in general. However, for the closed subspace  $l_0^\infty(\mathbb{Z}^d) \subset l^\infty(\mathbb{Z}^d)$  consisting of bounded sequences indexed by  $\mathbb{Z}^d$  vanishing at infinity,  $\lim_{x \rightarrow \infty} f(x) = 0$  for any  $f \in l_0^\infty(\mathbb{Z}^d)$ , the above proof immediately generalizes and yields the analogous compactness statement:  $(f_n)_{n \in \mathbb{N}}$  converges strongly in  $l_0^\infty(\mathbb{Z}^d)$  if and only if it converges weakly and

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \sup_{|x| > L} |f_n(x)| = 0.$$

## APPENDIX B. MULTI-LINEAR ESTIMATES.

In this section we gather some a-priori bounds for the multi-linear functional  $\mathcal{Q}_\mu^c$ , respectively  $\mathcal{Q}_\mu^d$ , which are used in the proof of Lemma 2.1, respectively Lemma 2.2. In the continuous case, the multi-linear bounds are an extension of our results in [24], in the discrete case the multi-linear bounds from [25] are much stronger than their corresponding continuous counterparts.

**B.1. Multi-linear estimates for  $\mathcal{Q}_\mu^c$ .** We extend, with some small simplifications, the proofs of the multi-linear estimates from [24] from the case  $\psi = \mathbf{1}_{[0,1]}$  to the more general densities needed here. Again, we will write  $\|f\|$  for the  $L^2$ -norm of a function  $f \in L^2(\mathbb{R})$  in this section. As soon as the four-linear functional  $L^2(\mathbb{R})^4 \ni (f_1, f_2, f_3, f_4) \mapsto \mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)$  is bounded, scaling  $f \mapsto f/\sqrt{\lambda}$ , shows that  $P_\lambda^c$  defined in (1.1) obeys

$$P_\lambda^c = P_1^c \lambda^2 \tag{B.1}$$

for all  $\lambda > 0$ . In particular, with  $\lambda = \|f\|^2$ ,

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq P_\lambda^c = P_1^c \|f\|^4. \tag{B.2}$$

For the boundedness we note

**Lemma B.1.** *Assume that  $\mu$  has a density  $\psi \in L^2(\mathbb{R})$ . Then  $P_1^c \leq 12^{-1/4} \|\psi\|$  and for any functions  $f_j \in L^2(\mathbb{R})$ ,  $j=1,2,3,4$ ,*

$$|\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| \leq P_1^c \prod_{j=1}^4 \|f_j\|. \tag{B.3}$$

Moreover, if  $0 \leq \psi \not\equiv 0$ , then  $P_1^c > 0$ .

*Proof.* We sketch the proof for the convenience of the reader. Using the triangle and generalized Hölder inequalities,

$$\begin{aligned} |\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| &\leq \iint_{\mathbb{R} \times \mathbb{R}} \prod_{j=1}^4 |T_t f_j| dx \psi(t) dt \leq \prod_{j=1}^4 \left( \iint_{\mathbb{R} \times \mathbb{R}} |T_t f_j|^4 dx |\psi(t)| dt \right)^{1/4} \\ &= \prod_{j=1}^4 (\mathcal{Q}_{|\psi|}^c(f_j, f_j, f_j, f_j))^{1/4}. \end{aligned}$$

Thus it is enough to show that  $\mathcal{Q}_{|\psi|}^c(f, f, f, f) \leq P_1^c \|f\|^4$  for some finite constant  $P_1^c$  and all  $f \in L^2(\mathbb{R})$ . Using the Cauchy-Schwarz inequality, one gets

$$\mathcal{Q}_\mu^c(f, f, f, f) = \iint_{\mathbb{R} \times \mathbb{R}} |T_t f|^{3+1} \psi(t) dx dt \leq \left( \iint_{\mathbb{R} \times \mathbb{R}} |T_t f|^6 dx dt \right)^{1/2} \left( \iint_{\mathbb{R} \times \mathbb{R}} |T_t f|^2 \psi(t)^2 dx dt \right)^{1/2}.$$

The second factor is seen to be bounded by  $\|\psi\| \|f\|$  doing the  $x$ -integration first, using that  $T_t$  is a unitary operator on  $L^2(\mathbb{R})$ . The first factor is bounded by the one-dimensional Strichartz inequality,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |T_t f(x)|^6 dx dt \leq S_1^6 \|f\|^6, \quad (\text{B.4})$$

which holds due to the dispersive properties of the free Schrödinger equation, [23, 50, 53]. The sharp constant in (B.4) is known,  $S_1 = 12^{-1/12}$ , one even knows  $S_2$  in two space dimensions, see [20, 26]. Thus

$$\mathcal{Q}_\mu^c(f, f, f, f) \leq S_1^3 \|\psi\| \|f\|^4.$$

Hence  $P_1^c \leq S_1^3 \|\psi\| = 12^{-1/4} \|\psi\|$ , using the sharp value for the Strichartz constant. This proves the upper bound on  $P_1^c$ .

To see that  $P_1^c > 0$ , note that  $P_1^c = 0$  would imply  $\int_{\mathbb{R}} \int_{\mathbb{R}} |T_t f(x)|^4 \psi(t) dx dt = 0$  for all  $f \in L^2(\mathbb{R})$ . Since  $\psi$  is non-negative,  $T_t f(x)$  has to be zero for Lebesgue almost every  $x$  and  $\psi(t) dt$  almost every  $t$ . Thus, by the unicity of  $T_t$ ,

$$0 = \int \int |T_t f(x)|^2 dx \psi(t) dt = \int \|T_t f\|^2 \psi(t) dt = \|f\|^2 \int \psi(t) dt.$$

Since  $\int \psi(t) dt > 0$  this shows  $\|f\| = 0$ . Hence  $P_1^c > 0$ . ■

**Remark B.2.** For a more explicit lower bound on  $P_1^c$  one can use a chirped Gaussian test-function similar to [30, 59], see also [24]. If the initial condition  $f$  is given by

$$f(x) = A_0 e^{-x^2/\sigma_0} \quad \text{with } \text{Re}(\sigma_0) > 0 \quad (\text{B.5})$$

then  $u(t, x) = A(t) e^{-x^2/\sigma(t)}$  solves the free Schrödinger equation if  $\sigma(t) = \sigma_0 + 4it$  and  $A(t) = A_0 \sqrt{\sigma_0}/\sqrt{\sigma(t)}$ . Hence

$$T_t f(x) = A(t) e^{-x^2/\sigma(t)}, \quad (\text{B.6})$$

for initial conditions of the form (B.5), see, e.g., [59]. Thus

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |T_t f|^4 dx \psi(t) dt = \sqrt{\frac{\pi}{4}} |A_0|^4 \frac{|\sigma_0|^2}{\sqrt{\text{Re} \sigma_0}} \int_{\mathbb{R}} \frac{1}{|\sigma(t)|} \psi(t) dt. \quad (\text{B.7})$$

Choosing  $|A_0|^2 = \sqrt{2 \text{Re}(\sigma_0)/(|\sigma_0|^2 \pi)}$  yields the normalization  $\|f\| = 1$  and hence

$$P_1^c \geq \sup_{\sigma_0 \in \mathbb{C}} \frac{\sqrt{\text{Re}(\sigma_0)}}{\sqrt{\pi}} \int_{\mathbb{R}} \frac{1}{\sqrt{\text{Re}(\sigma_0)^2 + (\text{Im}(\sigma_0) + 4t)^2}} \psi(t) dt > 0. \quad (\text{B.8})$$

if  $\psi$  is non-negative and positive on a set of positive Lebesgue measure.

The proof of the multi-linear estimates is based on the by now well-known bilinear Strichartz estimate, see, for example, [6, 13, 28, 43],

$$\|T_t f_1 T_t f_2\|_{L^2(\mathbb{R} \times \mathbb{R}, dt dx)} \lesssim \frac{1}{\sqrt{\text{dist}(\text{supp } \widehat{f}_1, \text{supp } \widehat{f}_2)}} \|f_1\|_{L^2(\mathbb{R})} \|f_2\|_{L^2(\mathbb{R})}. \quad (\text{B.9})$$

going back to [6]. For a simple explicit proof of (B.9), see for example, [31] or [24]. Now assume that the supports of  $\widehat{f}_l$  and  $\widehat{f}_m$  have positive distance for some  $l, m \in \{1, 2, 3, 4\}$ . Since

$$|\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| \leq \iint_{\mathbb{R} \times \mathbb{R}} \prod_{j=1}^4 |T_t f_j(x)| |\psi(t)| dx dt$$

we can assume that  $l = 1$  and  $m = 2$  without loss of generality. Then, by Cauchy-Schwarz followed by (B.9) and (B.3),

$$\begin{aligned} |\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| &\leq \|T_t f_1 T_t f_2\|_{L^2(\mathbb{R} \times \mathbb{R}, dt dx)} \left( \mathcal{Q}_{|\psi|^2}(f_3, f_3, f_4, f_4) \right)^{1/2} \\ &\lesssim \frac{\|\psi\|_{L^4}}{\sqrt{\text{dist}(\text{supp } \widehat{f}_1, \text{supp } \widehat{f}_2)}} \prod_{j=1}^4 \|f_j\|. \end{aligned}$$

This yields the first part of the following

**Lemma B.3** (Refined multi-linear estimates). *Let  $f_j \in L^2(\mathbb{R})$  for  $j = 1, 2, 3, 4$ .*

(i) *Let the measure  $\mu$  have density  $\psi \in L^4(\mathbb{R}, dt)$ . If, for some  $i \neq j$ , the supports of the Fourier transforms  $\widehat{f}_i$  and  $\widehat{f}_j$  are separated, i.e.,  $s = \text{dist}(\text{supp } \widehat{f}_i, \text{supp } \widehat{f}_j) > 0$ , then*

$$|\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| \lesssim \frac{1}{\sqrt{s}} \|f_1\| \|f_2\| \|f_3\| \|f_4\|. \quad (\text{B.10})$$

where the implicit constant depends only on  $\|\psi\|_{L^4(\mathbb{R}, dt)}$ .

(ii) *Let the measure  $\mu$  have density  $\psi \in L^4(\mathbb{R}, t^2 dt)$ . If, for some  $i \neq j$ , the supports of  $f_i$  and  $f_j$  are separated, i.e.,  $s = \text{dist}(\text{supp } f_i, \text{supp } f_j) > 0$ , then*

$$|\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| \lesssim \frac{1}{\sqrt{s}} \|f_1\| \|f_2\| \|f_3\| \|f_4\|. \quad (\text{B.11})$$

where the implicit constant in the bound depends only on  $\|\psi\|_{L^4(\mathbb{R}, t^2 dt)}$ .

It remains to show the second part of Lemma B.3. In fact, using a symmetry of  $\mathcal{Q}_\mu^c$  under Fourier transform, the second half of the Lemma is *equivalent* to the first half. In order to formulate this symmetry we need a little bit more notation: Given  $f \in L^2(\mathbb{R})$  let  $\check{f}$  be its inverse Fourier transform, that is, if  $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ , say,

$$\check{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ikx} f(x) dx. \quad (\text{B.12})$$

and given  $\psi \in L^2(\mathbb{R})$  define  $\widetilde{\psi}$  by

$$\widetilde{\psi}(\tau) = \frac{\psi(-1/(4\tau))}{2|\tau|}. \quad (\text{B.13})$$

A simple change of variables shows

$$\|\tilde{\psi}\| = \|\psi\|, \quad (\text{B.14})$$

i.e., the  $L^2$  norm is conserved.

**Lemma B.4.** (*Duality*) *Given a measure  $\mu$  with density  $\psi \in L^2(\mathbb{R})$  let the measure  $\tilde{\mu}$  have density  $\tilde{\psi}$ , where  $\tilde{\psi}$  is defined in (B.13). Then*

$$\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = \mathcal{Q}_{\tilde{\mu}}^c(\check{f}_1, \check{f}_2, \check{f}_3, \check{f}_4) \quad (\text{B.15})$$

for all  $f_j \in L^2(\mathbb{R})$ ,  $j=1,2,3,4$ .

We postpone the proof of Lemma B.4 for the moment. Given the duality (B.15), the equivalence of (B.11) and (B.10) is easy. For example, assume that the supports of  $f_1$  and  $f_2$  are separated by at least  $s$ , that is, the supports of the Fourier transforms of  $\check{f}_1$  and  $\check{f}_2$  are separated by at least  $s$ . Hence (B.15) in tandem with (B.10) yields

$$|\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)| = |\mathcal{Q}_{\tilde{\mu}}^c(\check{f}_1, \check{f}_2, \check{f}_3, \check{f}_4)| \lesssim \frac{1}{\sqrt{s}} \prod_{j=1}^4 \|\check{f}_j\| = \frac{1}{\sqrt{s}} \prod_{j=1}^4 \|f_j\|$$

where the implicit constant in the inequality depends only on  $\|\tilde{\psi}\|_{L^4(\mathbb{R}, dt)}$ . Since  $\|\tilde{\psi}\|_{L^4(\mathbb{R}, dt)} = 4^{1/4} \|\psi\|_{L^4(\mathbb{R}, t^2 dt)}$  this proves (B.11). It remains to give the

*Proof of Lemma B.4.* The duality is a simple consequence of the so-called quasi-conformal symmetry of the free Schrödinger evolution, which is true in all space dimensions, but we need only the one-dimensional case here. The solution  $u(t, x) = T_t f(x)$  of  $i\partial_t u = -\partial_x^2 u$  with initial condition  $u(0, x) = f(x)$  is given by

$$u(t, x) = \frac{1}{\sqrt{4\pi it}} \int_{\mathbb{R}} e^{i\frac{|x-y|^2}{4t}} f(y) dy \quad (\text{B.16})$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixk} e^{-itk^2} \hat{f}(k) dk. \quad (\text{B.17})$$

Expanding the square in (B.16) gives

$$u(t, x) = \frac{1}{\sqrt{2it}} e^{i\frac{x^2}{4t}} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\frac{xy}{2t}} e^{i\frac{y^2}{4t}} f(y) dy$$

and comparing this with (B.17), one sees

$$u(t, x) = \frac{1}{\sqrt{2it}} e^{i\frac{x^2}{4t}} \tilde{u}\left(-\frac{1}{4t}, -\frac{x}{2t}\right) \quad (\text{B.18})$$

where  $\tilde{u}(\tau, y) = T_\tau \check{f}(y)$  is the solution of the free Schrödinger equation with initial condition  $\check{f}$ . Hence, setting  $u_j(t, x) = T_t f_j(x)$ ,  $\tilde{u}_j(\tau, y) = T_\tau \check{f}_j(y)$  and using (B.18) one sees

$$\begin{aligned} \mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \overline{u_1(t, x)} u_2(t, x) \overline{u_3(t, x)} u_4(t, x) dx \psi(t) dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \overline{\tilde{u}_1(\tau, y)} \tilde{u}_2(\tau, y) \overline{\tilde{u}_3(\tau, y)} \tilde{u}_4(\tau, y) dy \frac{\psi(-1/(4\tau))}{2|\tau|} d\tau \\ &= \mathcal{Q}_{\tilde{\mu}}^c(\check{f}_1, \check{f}_2, \check{f}_3, \check{f}_4) \end{aligned} \quad (\text{B.19})$$

where we used (B.18), did the change of variables  $x = 2ty$ ,  $dx = 2|t|dy$ , and then  $t = -1/(4\tau)$ ,  $dt = d\tau/(4\tau^2)$ , and defined  $\tilde{\psi}$  by (B.13). This proves (B.15).  $\blacksquare$

As a last tool, we note that although  $\mathcal{Q}_\mu^c$  is certainly not local, it is ‘nearly’ local in the following sense.

**Lemma B.5** (Quasi-locality). *Assume that the measure  $\mu$  has density  $\psi \in L^2(\mathbb{R})$ . If, for some  $s > 0$ ,  $f_1$  has support outside the interval  $[-3s, 3s]$  and for  $j = 2, 3, 4$  all the functions  $f_j$  have support within  $[-s, s]$ , then*

$$\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = 0 \quad (\text{B.20})$$

*The same is true if  $\hat{f}_1$  has support outside the interval  $[-3s, 3s]$  and for  $j = 2, 3, 4$  all the functions  $\hat{f}_j$  have support within  $[-s, s]$*

*Proof.* As in [24], one uses the explicit representation (B.16) for the free time evolution to see that

$$\begin{aligned} \mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) &= \int_{\mathbb{R}} \frac{\psi(t)dt}{(4\pi t)^2} \int_{\mathbb{R}} dx \int_{\mathbb{R}^4} e^{\frac{ix(y_1 - y_2 + y_3 - y_4)}{2t}} e^{\frac{-i(y_1^2 - y_2^2 + y_3^2 - y_4^2)}{4t}} \overline{f_1(y_1)} f_2(y_2) \overline{f_3(y_3)} f_4(y_4) dy \\ &= \frac{1}{8\pi^2} \int_{\mathbb{R}} \frac{\psi(t)dt}{|t|} \int_{\mathbb{R}} dz \int_{\mathbb{R}^4} e^{i(y_1 - y_2 + y_3 - y_4)z} e^{\frac{-i(y_1^2 - y_2^2 + y_3^2 - y_4^2)}{4t}} \overline{f_1(y_1)} f_2(y_2) \overline{f_3(y_3)} f_4(y_4) dy \\ &= \frac{1}{4\pi} \int_{\mathbb{R}} \frac{\psi(t)dt}{|t|} \int_{\mathbb{R}^4} \delta(y_1 - y_2 + y_3 - y_4) e^{\frac{-i(y_1^2 - y_2^2 + y_3^2 - y_4^2)}{4t}} \overline{f_1(y_1)} f_2(y_2) \overline{f_3(y_3)} f_4(y_4) dy \end{aligned} \quad (\text{B.21})$$

where we first made the change of variables  $x = 2tz$ ,  $dx = 2|t|dz$  and then used  $\frac{1}{2\pi} \int_{\mathbb{R}} e^{iz\eta} dz = \delta(\eta)$ , as distributions. Since the  $y$ -integration in formula (B.21) is restricted to the 3 dimensional subspace given by  $0 = y_1 - y_2 + y_3 - y_4$  one has  $\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = 0$  whenever  $f_1(y_1)f_2(y_2)f_3(y_3)f_4(y_4) = 0$  on this subspace. This proves the first assertion.

For the Fourier-space version, i.e., under the conditions stated on the Fourier transforms of  $f_j$ , simply note that the duality Lemma B.4 says

$$\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = \mathcal{Q}_\mu^c(\check{f}_1, \check{f}_2, \check{f}_3, \check{f}_4)$$

Thus (B.21) applied to  $\mathcal{Q}_\mu^c$  shows that  $\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = 0$  whenever the product  $\check{f}_1(k_1)\check{f}_2(k_2)\check{f}_3(k_3)\check{f}_4(k_4) = 0$  on the subspace given by  $k_1 - k_2 + k_3 - k_4 = 0$ . Since  $\check{f}(k) = \hat{f}(-k)$ , this proves the second assertion of the Lemma.  $\blacksquare$

**Remarks B.6.** (i) The proof of Lemma B.5 can be interpreted as a non-resonance effect: If the four wave-packets  $f_1, f_2, f_3, f_4$  are *non-resonant* in the sense that  $\text{supp}(f_1) - \text{supp}(f_2) + \text{supp}(f_3) - \text{supp}(f_4) = 0$  or  $\text{supp}(\hat{f}_1) - \text{supp}(\hat{f}_2) + \text{supp}(\hat{f}_3) - \text{supp}(\hat{f}_4) = 0$  then  $\mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4) = 0$ .

(ii) Although the quasi-locality is not needed for our existence proof of maximizers, it is needed in the proof of regularity of weak solution of the dispersion management equation (1.22). Given Lemmata B.1, B.3, and B.5, a straightforward adaptation of the proof in [24], which in our notation, was given there for the special case  $\psi = \mathbf{1}_{[0,1]}$ , shows



that all weak solutions of the Gabitov-Turitsyn equation for vanishing average dispersion (1.22) are already Schwartz functions as soon as  $\psi \in L^2(\mathbb{R}) \cap L^4(\mathbb{R}) \cap L^4(\mathbb{R}, t^2 dt)$ .

**B.2. Multi-linear estimates for  $\mathcal{Q}_\mu^d$ .** In this section we will use  $\|f\|$  for the  $l^2$ -norm of a sequence  $f \in l^2(\mathbb{Z})$ . Recall that  $P_\lambda^d$  is defined in (1.2). As in the continuous case, a simple scaling argument shows

$$P_\lambda^d = P_1^d \lambda^2 \quad (\text{B.22})$$

for all  $\lambda > 0$ . Thus, with  $\lambda = \|f\|^2$ ,

$$\mathcal{Q}_\mu^d(f, f, f, f) \leq P_\lambda^d = P_1^d \|f\|^4. \quad (\text{B.23})$$

**Lemma B.7.** *Let  $\mu$  be a bounded measure. For any  $f_j \in l^2(\mathbb{Z}^d)$ ,  $j = 1, 2, 3, 4$ , we have*

$$|\mathcal{Q}_\mu^d(f_1, f_2, f_3, f_4)| \leq \mu(\mathbb{R}) \prod_{j=1}^4 \|f_j\|, \quad (\text{B.24})$$

in particular,  $0 < P_\lambda^d \leq \mu(\mathbb{R}) \lambda^2$  for any  $\lambda > 0$ .

*Proof.* By the triangle and Hölder inequalities

$$|\mathcal{Q}_\mu^d(f_1, f_2, f_3, f_4)| \leq \int_{\mathbb{R}} \sum_{x \in \mathbb{Z}^d} \prod_{j=1}^4 |S_t f_j(x)| \mu(dt) \leq \int_{\mathbb{R}} \prod_{j=1}^4 \|S_t f_j\|_{l^4} \mu(dt)$$

On the other hand, on the scale of  $l^p$ -spaces the simple but strong inequality  $\|g\|_{l^q} \leq \|g\|_{l^p}$  holds for all  $1 \leq p \leq q \leq \infty$ . So  $\|S_t f_j\|_{l^4} \leq \|S_t f_j\|_{l^2} = \|f_j\|_{l^2}$  since  $S_t = e^{it\Delta}$  is unitary on  $l^2(\mathbb{Z})$ . The proof of  $0 < P_\lambda^d$  is similar to the continuous case.  $\blacksquare$

The following refined multi-linear estimate for  $\mathcal{Q}_\mu^d$  is from [25] where also the multi-dimensional case is done. Note that the refined multi-linear estimate for  $\mathcal{Q}_\mu^d$  shows a stronger decay than their continuous counterparts and as Lemma B.7 this decay holds under the weakest possible assumption on  $\mu$ .

**Lemma B.8.** *Let  $\mu$  be a bounded measure with bounded support and assume that  $s = \text{dist}(\text{supp}(f_k), \text{supp}(f_l)) > 0$  for some  $j, k \in \{1, 2, 3, 4\}$ . Then*

$$|\mathcal{Q}_\mu^d(f_1, f_2, f_3, f_4)| \lesssim |s|^{-\delta|s|} \prod_{j=1}^4 \|f_j\| \quad (\text{B.25})$$

for any  $0 < \delta < 1/2$ , where the implicit constant depends only on  $\delta$ .

*Sketch of proof of Lemma B.8:* The proof of Lemma B.8 rests on the strong bilinear bound

$$\sup_{t \in [-\tau, \tau]} \|(e^{it\Delta} f_1)(e^{it\Delta} f_2)\| \lesssim s^{-\delta s} \|f_1\| \|f_2\|. \quad (\text{B.26})$$

for all  $0 < \delta < 1/2$ , with  $e^{it\Delta}$  the free discrete one-dimensional Schrödinger evolution and  $s = \text{dist}(\text{supp}(f_1), \text{supp}(f_2))$ . Once one has (B.26) the bound (B.25) follows as in the continuous case.

The estimate (B.26) itself follows from the bound

$$\sup_{t \in [-\tau, \tau]} |\langle x | e^{it\Delta} | y \rangle| \leq \min(1, e^{4\tau} \frac{(4\tau)^{|x-y|}}{|x-y|!}) \quad (\text{B.27})$$

for the kernel of the free time evolution  $e^{it\Delta}$ ,  $x, y \in \mathbb{Z}$  and  $0 \leq \tau < \infty$ . Here  $\langle x | M | y \rangle = \langle \delta_x, M \delta_y \rangle$  for an operator  $M$  on  $l^2(\mathbb{Z})$ , where  $\delta_x$  is the Kronecker delta-function. The bound (B.27) shows that unlike to the continuous case, the kernel of the free discrete Schrödinger evolution has a *strong point-wise decay* locally uniformly in  $t$ . This is due to the finite speed of propagation for the discrete Schrödinger equation, or, in other words, the Fourier spectrum of the lattice  $\mathbb{Z}$  is the bounded interval  $[-\pi, \pi]$ . The easiest way to see (B.27) is to note that since the discrete one-dimensional Laplace is bounded with norm  $\|\Delta\| = 4$ , the free discrete Schrödinger evolution can be written as a norm-converging exponential series  $e^{it\Delta} = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \Delta^n$ . Thus

$$|\langle x | e^{it\Delta} | y \rangle| \leq \sum_{n=0}^{\infty} \frac{(|t|)^n}{n!} |\langle x | \Delta^n | y \rangle| = \sum_{n=|x-y|}^{\infty} \frac{(4|t|)^n}{n!} \leq e^{4\tau} \frac{(4\tau)^{|x-y|}}{|x-y|!}$$

since  $\langle x | \Delta^n | y \rangle = 0$  if  $|x-y| > n$  and  $|\langle x | \Delta^n | y \rangle| \leq \|\Delta\|^n \leq 4^n$ . For more details and extensions to  $l^2(\mathbb{Z}^d)$  with  $d > 1$ , see [25]. ■

### APPENDIX C. SHIFTS, BOOSTS, AND GALILEI TRANSFORMATIONS.

We will only discuss the one-dimensional case which is somewhat easier than Galilei transformations on  $L^2(\mathbb{R}^d)$  since we do not have to deal with rotations in one dimension. The unitary operator implementing the shift  $S_y \xi : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ ,  $(S_y f)(x) = f(x-y)$  is given by

$$S_\xi = e^{-i\xi P} \quad (\text{C.1})$$

where  $P = -i\partial_x$  is the momentum operator. Indeed, since  $e^{-i\xi P}$  corresponds to multiplication by  $e^{-i\xi k}$  in Fourier space, we have

$$(e^{-i\xi P} f)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i(x-\xi)k} \widehat{f}(k) dk = f(x-\xi).$$

Boosts, i.e., shifts in momentum space are given by  $e^{iv\cdot} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ , i.e., multiplication by  $e^{ivx}$ , since

$$\widehat{e^{iv\cdot} f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix(k-v)} f(x) dx = \widehat{f}(k-v). \quad (\text{C.2})$$

Finally, if  $G$  is a bounded (measurable) function then  $G(P)$  is defined by

$$\widehat{G(P)f}(k) = G(k)\widehat{f}(k).$$

Of course, for any  $\xi \in \mathbb{R}$  the operators  $G(P)$  and  $e^{-i\xi P}$  commute,  $G(P)e^{-i\xi P} = e^{-i\xi P}G(P)$ . Moreover, for any  $v \in \mathbb{R}$  the commutation relation

$$G(P)e^{iv\cdot} = e^{iv\cdot}G(P+v) \quad (\text{C.3})$$

holds. Indeed, Computing the Fourier transform yields

$$\begin{aligned} (G(P)e^{iv\cdot}f)\widehat{\phantom{x}}(k) &= G(k)\widehat{e^{iv\cdot}f}(k) = G(k)\widehat{f}(k-v) \\ &= (G(\cdot+v)\widehat{f})(k-v) = (G(P+v)f)\widehat{\phantom{x}}(k-v) \\ &= (e^{iv\cdot}G(P+v)f)\widehat{\phantom{x}}(k). \end{aligned}$$

In particular, choosing  $G(P) = e^{-itP^2}$ , we see the commutation relation

$$\begin{aligned} e^{-itP^2}e^{iv\cdot}e^{-i\xi P} &= e^{iv\cdot}e^{-i\xi P}e^{-it(P+v)^2} = e^{iv\cdot}e^{-i\xi P}e^{-it(P^2+2vP+v^2)} \\ &= e^{-itv^2}e^{iv\cdot}e^{-i(\xi+2tv)P}e^{-itP^2}. \end{aligned} \tag{C.4}$$

Now let  $f \in L^2(\mathbb{R})$ . Then  $u(t) = T_t f = e^{-itP^2}f$  is the solution of the (one-dimensional) Schrödinger equation  $-i\partial_t u = P^2 u = -\partial_x^2 u$  with initial condition  $u(0) = f$ . Using (C.4), the solution of the free Schrödinger equation for the translated and boosted initial condition  $f_{\xi,v} = e^{iv\cdot}e^{-i\xi P}f$  is given by

$$\begin{aligned} u_{\xi,v}(t, x) &:= e^{-itP^2}f_{\xi,v}(t, x) = (e^{-itP^2}e^{iv\cdot}e^{-i\xi P}f)(x) \\ &= (e^{-itv^2}e^{iv\cdot}e^{-i(\xi+2tv)P}e^{-itP^2}f)(x) \\ &= e^{-itv^2}e^{ivx}(e^{-i(\xi+2tv)P}e^{-itP^2}f)(x) \\ &= e^{-itv^2}e^{ivx}(e^{-itP^2}f)(x - \xi - 2tv) \\ &= e^{-itv^2}e^{ivx}u(t, x - \xi - 2tv), \end{aligned} \tag{C.5}$$

that is, on the level of the solutions of the free time-dependent Schrödinger equation, translations and boost of the initial condition are implemented by the Galilei transformations  $\mathcal{G}_{\xi,v}$  given by  $(\mathcal{G}_{\xi,v}u)(t, x) := u_{\xi,v}(t, x) = e^{-itv^2}e^{ivx}u(t, x - \xi - 2tv)$ . Except for the time-dependent phase factor  $e^{-itv^2}$ , formula (C.5) is exactly what one would have guessed from classical mechanics

Note that  $P^2 = -\Delta$ . A simple calculation now shows that the functional

$$f \mapsto \mathcal{Q}_\mu^c(f, f, f, f) = \int_{\mathbb{R}} \int_{\mathbb{R}} |(e^{-itP^2}f)(x)|^4 dx \mu(dt)$$

is invariant under translations and boosts in  $L^2(\mathbb{R})$ . Similarly, it is straightforward to see that the 4-linear functional  $f_j \mapsto \mathcal{Q}_\mu^c(f_1, f_2, f_3, f_4)$  is invariant under simultaneous shifts and boosts of the  $f_j$ .

**Acknowledgements:** It is a pleasure to thank Vadim Zharnitsky for instructive discussions on the dispersion management technique. Dirk Hundertmark thanks the Max-Planck Institute for Physics of Complex Systems in Dresden and the Max-Planck Institute for Mathematics in the Sciences in Leipzig for their warm hospitality while part of this work was done.

## REFERENCES

- [1] M.J. Ablowitz and G. Biondini, *Multiscale pulse dynamics in communication systems with strong dispersion management*. Opt. Lett. **23** (1998), 1668–1670.

- [2] M. Ablowitz and Z. H. Musslimani, *Discrete Diffraction Managed Spatial Solitons*. Phys. Rev. Lett. **87** (2001), 254102 [4 pages].
- [3] M. Ablowitz and Z. H. Musslimani, *Discrete vector spatial solitons in a nonlinear waveguide array*. Phys. Rev. E **65** (2002), 056618.
- [4] M. Ablowitz and Z. H. Musslimani, *Discrete spatial solitons in a diffraction managed nonlinear waveguide array: a unified approach*. Physica D **184** (2003), 276-303.
- [5] M. Sh. Birman and M. Z. Solomjak, *Spectral theory of selfadjoint operators in Hilbert space*. Translated from the 1980 Russian original by S. Khrushchv and V. Peller. Mathematics and its Applications (Soviet Series). D. Reidel Publishing Co., Dordrecht, 1987. xv+301 pp.
- [6] J. Bourgain, *Global solutions of nonlinear Schrödinger equations*. American Mathematical Society Colloquium Publications, **46**. AMS, Providence, RI, 1999.
- [7] D. Cai, A. R. Bishop, and N. Groenbech-Jensen, *Localized states in discrete nonlinear Schrödinger equations*. Phys. Rev. Lett. **72** (1994), 591-595.
- [8] Thierry Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics **10**, AMS, Providence, Rhode Island, 2003.
- [9] T. Cazenave and P.-L. Lions, *Orbital stability of standing waves for some nonlinear Schrödinger equations*, Comm. Math. Phys. **85** (1982), no. 4, 549-561.
- [10] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, R. M. Derosier, *8 10 Gb/s transmission through 280 km of dispersion-managed fiber*. IEEE Phot. Tech. Lett. **5** (1993), 1233-1235.
- [11] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, R. M. Derosier, E. R. Giles, B. M. Nyman, G. A. Ferguson, J. W. Sulhoff, J. L. Zyskind, *One-third terabit/s transmission through 150 km of dispersion-managed fiber*, Photonics Technology Letters, IEEE (1995), **7**, Issue: 1, 98-100.
- [12] D. N. Christodoulides and R. I. Joseph, *Discrete self-focusing in nonlinear arrays of coupled waveguides*. Opt. Lett. **13** (1988), 794-796.
- [13] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, *A refined global well-posedness result for Schrödinger equations with derivative*. SIAM J. Math. Anal. **34** (2002), 64-86.
- [14] A. S. Davydov, *Theory of contraction of proteins under their excitation*. J. Theor. Biol. **38** (1973), 559-569.
- [15] H. Eisenberg, Y. Silverberg, R. Morandotti, A. Boyd, and J. Aitchison, J. (1998). *Discrete spatial optical solitons in waveguide arrays*. Phys. Rev. Lett. **81** (1998), 3383-3386.
- [16] H. Eisenberg, Y. Silverberg, R. Morandotti, and J. Aitchison, *Diffraction management*. Phys. Rev. Lett. **85** (2000), 1863-1866.
- [17] I. Ekeland, *Sur les problèmes variationnels*. (French) C. R. Acad. Sci. Paris Sér. A-B **275** (1972), A1057-A1059.
- [18] I. Ekeland, *Remarques sur les problèmes variationnels*. (French) C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1347-A1348.
- [19] M. B. Erdoğan, D. Hundertmark, and Y.-R. Lee, *Exponential decay of dispersion managed solitons for vanishing average dispersion*, preprint.
- [20] Damiano Foschi, *Maximizers for the Strichartz inequality*. J. Eur. Math. Soc. **9** (2007), 739-774.
- [21] I. Gabitov and S.K. Turitsyn, *Averaged pulse dynamics in a cascaded transmission system with passive dispersion compensation*. Opt. Lett. **21** (1996), 327-329.
- [22] I. Gabitov and S.K. Turitsyn, *Breathing solitons in optical fiber links*. JETP Lett. **63** (1996) 861.

- [23] Jean Ginibre and Giorgio Velo, *The global Cauchy problem for the nonlinear Schrödinger equation*. Ann. Inst. H. Poincaré Anal. Non Linéaire **2** (1985), 309–327.
- [24] D. Hundertmark, and Y.-R. Lee, *Decay estimates and smoothness for solutions of the dispersion managed non-linear Schrödinger equation*. Comm. Math. Phys. **286** (2009), no. 3, 851–873.
- [25] D. Hundertmark, and Y.-R. Lee, *Super-exponential decay of diffraction management solitons*, preprint.
- [26] Dirk Hundertmark and Vadim Zharnitsky, *On sharp Strichartz inequalities for low dimensions*. International Mathematics Research Notices, vol. 2006, Article ID 34080, 18 pages, 2006. doi:10.1155/IMRN/2006/34080
- [27] Y. Jabri, *The mountain pass theorem. Variants, generalizations, and some applications*. Encyclopedia of Mathematics and its Applications **95**.
- [28] C.E. Kenig, G. Ponce, and L. Vega, *Oscillatory integrals and regularity of dispersive equations*. Indiana Univ. Math. J. **40** (1991), 33–68.
- [29] S. Kumar and A. Hasegawa, *Quasi-soliton propagation in dispersion-managed optical fibers*. Opt. Lett. **22** (1997), 372–374.
- [30] Markus Kunze, *On a variational problem with lack of compactness related to the Strichartz inequality*. Calc. Var. Partial Differential Equations **19** (2004), no. 3, 307–336.
- [31] M. Kunze, J. Moeser, and V. Zharnitsky, *Ground states for the higher-order dispersion managed NLS equation in the absence of average dispersion*, J. Differential Equations **209** (2005), no. 1, 77–100.
- [32] Kurtzke, *Suppression of fiber nonlinearities by appropriate dispersion management*. IEEE Phot. Tech. Lett. **5** (1993), 1250–1253.
- [33] T. Lakoba and D.J. Kaup, *Shape of the stationary pulse in the strong dispersion management regime*. Electron. Lett. **34** (1998), 1124–1125.
- [34] L.D. Landau and E.M. Lifshitz, *Course of theoretical physics. Vol. 1. Mechanics*. Third edition. Pergamon Press, Oxford-New York-Toronto, Ont., 1976.
- [35] E. H. Lieb and M. Loss, *Analysis*. Second edition. Graduate Studies in Mathematics, **14**. AMS, Providence, RI, 2001.
- [36] C. Lin, H. Kogelnik, and L. G. Cohen, *Optical pulse equalization and low dispersion transmission in singlemode fibers in the 1.3–1.7  $\mu\text{m}$  spectral region*. Opt. Lett. **5** (1980), 476–478.
- [37] P. L. Lions, *The concentration-compactness principle in the calculus of variations. The locally compact case, part 1 and 2*. Annales de l’institut Henri Poincaré (C) Analyse non linéaire **1** no. 2 and no. 4 (1984), 109–145 and 223–283
- [38] Pavel M. Lushnikov, *Oscillating tails of dispersion-managed soliton*. J. Opt. Soc. Am. B **21** (2004), 1913–1918.
- [39] P.V. Mamyshev and N.A. Mamysheva, *Pulseoverlapped dispersion-managed data transmission and intrachannel four-wave mixing*. Opt. Lett. **24** (1999) 1454–1456.
- [40] Jamison Moeser, *Diffraction managed solitons: asymptotic validity and excitation thresholds* Nonlinearity **18** (2005), 2275–2297.
- [41] L.F. Mollenauer, A. Grant, X. Liu, X. Wei, C. Xie, and I. Kang, *Experimental test of dense wavelengthdivision multiplexing using novel, periodic-group-delaycomplemented dispersion compensation and dispersionmanaged solitons*. Opt. Lett. **28** (2003), 2043–2045.
- [42] L.F. Mollenauer, P.V. Mamyshev, J. Gripp, M.J. Neubelt, N. Mamysheva, L. Grüner-Nielsen, and T. Veng, *Demonstration of massive wavelength-division multiplexing over transoceanic distances by use of dispersionmanaged solitons*. Opt. Lett. **25** (1999), 704–706.

- [43] T. Ozawa and Y. Tsutsumi, *Space-time estimates for null gauge forms and nonlinear Schrödinger equations*. Differential Integral Equations **11** (1998), 201–222.
- [44] P. Panayotaros, *Breather solutions in the diffraction managed NLS equation*. Physica D **206** (2005), 213–231.
- [45] M. Reed and B. Simon, *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press, New York–London, 1978.
- [46] A. C. Scott, 1985 *Davydov solitons in polypeptides* Phil. Trans. R. Soc. Lond. A **315** (1985), 423–436.
- [47] B. Simon, *Trace ideals and their applications*. Second edition. Mathematical Surveys and Monographs, **120**. AMS, Providence, RI, 2005. viii+150 pp.
- [48] Milena Stanislavova, *Regularity of ground state solutions of DMNLS equations*. J. Diff. Eq. **210** (2005), no. 1, 87–105.
- [49] Milena Stanislavova, *Diffraction Managed Solitons with Zero Mean Diffraction* Journal of Dynamics and Differential Equations **19** (2007), no. 2., 295–307.
- [50] R. S. Strichartz, *Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations*, Duke Math. J. **44** (1977), 705–714.
- [51] R. S. Strichartz, *The Way of Analysis*. Revised Edition. Jones and Bartlett publications, (2000).
- [52] A. Sukhorukov, Y. Kivshar, E. H. Eisenberg, and Y. Silberberg, *Spatial optical solitons in waveguide arrays*. IEEE J. Quantum Electron. **39** (2003), 31–50.
- [53] Cathrine Sulem and Pierre-Louis Sulem, *The non-linear Schrödinger equation. Self-focusing and wave collapse*. Applied Mathematical Sciences, **139**. Springer-Verlag, New York, 1999.
- [54] A. Trombettoni and A. Smerzi, *Discrete solitons and breathers with dilute BoseEinstein condensates*. Phys. Rev. Lett. **86** (1998), 2353–2356.
- [55] S.K. Turitsyn, N.J. Doran, J.H.B. Nijhof, V.K. Mezentsev, T. Schäfer, and W. Forysiak, in *Optical Solitons: Theoretical challenges and industrial perspectives*, V.E. Zakharov and S. Wabnitz, eds. (Springer Verlag, Berlin, 1999), p. 91.
- [56] S. K. Turitsyn, E. G. Shapiro, S. B. Medvedev, M. P. Fedoruk, and V. K. Mezentsev, *Physics and mathematics of dispersion-managed optical solitons*, Comptes Rendus Physique, Académie des sciences/Éditions scientifiques et médicales **4** (2003), 145–161.
- [57] M. I. Weinstein, *Lyapunov stability of ground states of nonlinear dispersive evolution equations*, Comm. Pure Appl. Math. **39** (1986), no. 1, 51–67.
- [58] M. I. Weinstein, *Excitation thresholds for nonlinear localized modes on lattices*, Nonlinearity **12** (1999), 673–91.
- [59] V. Zharnitsky, E. Grenier, C.K.R.T. Jones, and S.K. Turitsyn, *Stabilizing effects of dispersion management*, Physica D **152-153** (2001), 794–817.

DEPARTMENT OF MATHEMATICS, ALTGELD HALL, AND INSTITUTE FOR CONDENSED MATTER THEORY AT THE UNIVERSITY OF ILLINOIS AT URBANA–CHAMPAIGN, 1409 W. GREEN STREET, URBANA, IL 61801.

*E-mail address:* dirk@math.uiuc.edu

DEPARTMENT OF MATHEMATICS, SOGANG UNIVERSITY, SHINSU-DONG 1, MAPO-GU, SEOUL, 121-742, SOUTH KOREA.

*E-mail address:* younglee@sogang.ac.kr